

## A NEW CATALOGUE OF TYPE 1 AGN AND ITS IMPLICATION ON THE AGN UNIFIED MODEL

KYUSEOK OH<sup>1</sup>, SUKYOUNG K. YI<sup>2,3</sup>, KEVIN SCHAWINSKI<sup>1</sup>, MICHAEL KOSS<sup>1,†</sup>, BENNY TRAKHTENBROT<sup>1,‡</sup>*Draft version April 29, 2015*

## ABSTRACT

We have newly identified a substantial number of type 1 active galactic nuclei (AGN) featuring weak broad-line regions (BLRs) at  $z < 0.2$  from detailed analysis of galaxy spectra in the Sloan Digital Sky Survey Data Release 7. These objects predominantly show a stellar continuum but also a broad H $\alpha$  emission line, indicating the presence of a low-luminosity AGN oriented so that we are viewing the central engine directly without significant obscuration. These accreting black holes have previously eluded detection due to their weak nature. The new BLR AGNs we found increased the number of known type 1 AGNs by 49%. Some of these new BLR AGNs were detected at the *Chandra* X-ray Observatory, and their X-ray properties confirm that they are indeed type 1 AGN. Based on our new and more complete catalogue of type 1 AGNs, we derived the type 1 fraction of AGNs as a function of [O III]  $\lambda 5007$  emission luminosity and explored the possible dilution effect on the obscured AGN due to star-formation. The new type 1 AGN fraction shows much more complex behavior with respect to black hole mass and bolometric luminosity than suggested by the existing receding torus model. The type 1 AGN fraction is sensitive to both of these factors, and there seems to be a sweet spot (ridge) in the diagram of black hole mass and bolometric luminosity. Furthermore, we present a hint that the Eddington ratio plays a role in determining the opening angles.

*Subject headings:* galaxies: active — galaxies: Seyfert — galaxies: nuclei — quasars: general — galaxies: statistics — methods: data analysis

## 1. INTRODUCTION

According to the unification scheme of active galactic nuclei (AGNs, Antonucci 1993; Urry & Padovani 1995), which is widely accepted in current astrophysics, different viewing angles explain the variety of phenomenological subclasses of AGNs. For example, the optical spectra of type 1 AGNs present broad permitted emission lines as well as narrow emission lines exposing their central nuclear region, whereas type 2 AGNs show only narrow emission lines. According to the simplest AGN unification model and its strictest interpretation, the feasibility of observing broad permitted emission lines depends solely on the angle between the observer's line of sight and the axis of an obscuring medium, often called the “dust torus”.

With the aid of mega-scale surveys such as the Sloan Digital Sky Survey (SDSS; York et al. 2000) that provide homogenous photometric and spectroscopic data, AGNs have become one of the most actively studied topics in astrophysics. In particular, studies on *obscured AGNs* showing narrow emission lines have shed light on the connection between the host galaxies and the central AGN (Kauffmann et al. 2003; Salim et al. 2007; Schawinski et al. 2007; Westoby et al. 2007; Schawinski et al. 2009; Haggard et al. 2010; Schawinski et al. 2010).

*Unobscured* (type 1) AGNs reveal their bright nuclear light directly, and it has been an important task and

subject of interest to decompose the observed light into a central nuclear component and a diffuse galaxy component. For example, the high-spatial-resolution Advanced Camera for Surveys on the *Hubble Space Telescope*, has been extensively used for this purpose (Sánchez et al. 2004; Ballo et al. 2007; Alonso-Herrero et al. 2008; Simmons & Urry 2008; Schawinski et al. 2011; Simmons et al. 2013) in combination with two-dimensional surface brightness fitting techniques (Peng et al. 2002).

The fraction of type 1 AGN (among all AGNs) can act as an effective test of any model for AGN classifications and can be used to extract and constrain critical information, such as the opening angle and source luminosity, in the assumed framework. Based on a sample of high-luminosity and radio-selected broad-line AGNs, Willott et al. (2000) found the type 1 AGN fraction of 40% and a typical value of the “half-opening” angle of about 53°. The precise value of the type 1 AGN fraction has been reported to be substantially lower than this finding, when low-redshift surveys of lower optical luminosity cuts are used instead (Osterbrock & Shaw 1988; Huchra & Burg 1992; Maiolino & Rieke 1995; Maia et al. 2003). This difference seems to be sensible, as more recent, multiwavelength studies found a strong, positive luminosity dependence of the type 1 AGN fraction (Steffen et al. 2003; Hao et al. 2005; Simpson 2005; Hasinger 2008; Burlon et al. 2011; Assef et al. 2013).

AGNs exhibit characteristic spectral features across the electromagnetic spectrum: e.g., [O III]  $\lambda 5007$  optical emission lines, mid-infrared and hard X-ray luminosity as isotropic quantities tracing the source luminosity, which are naturally employed by many type 1 AGN fraction studies. In particular, Simpson (2005) found that the type 1 AGN fraction gradually increases with  $L[\text{O III}]$  based on SDSS Data Release 2 (Abazajian et al. 2004)

<sup>1</sup> Institute for Astronomy, Department of Physics, ETH Zurich, Wolfgang-Pauli-Strasse 27, CH-8093 Zurich, Switzerland; kyuseok.oh@phys.ethz.ch

<sup>2</sup> Department of Astronomy, Yonsei University, Seoul 120-749, Republic of Korea

<sup>3</sup> Yonsei University Observatory, Yonsei University, Seoul 120-749, Republic of Korea; yi@yonsei.ac.kr

<sup>†</sup> Ambizione fellow

<sup>‡</sup> Zwicky fellow

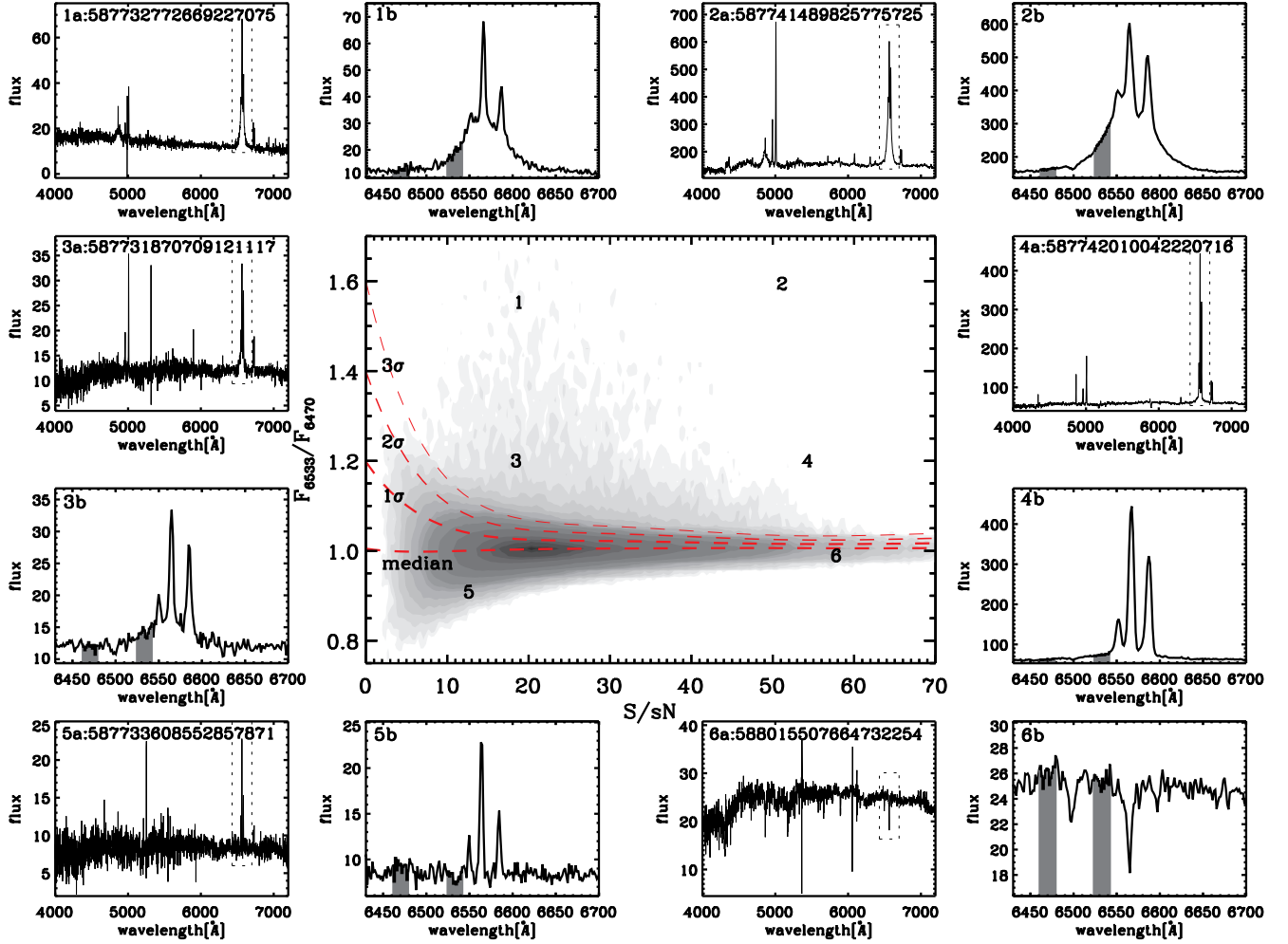


FIG. 1.— Flux ratio ( $F_{6,533}/F_{6,470}$ ) distribution versus signal-to-statistical noise ( $S/SN$ ). Demarcation lines in the central panel are shown as red dashed lines. The three lines indicate 84.1%, 97.7%, and 99.8% distributions. Example spectra represented by the numbered markings in the central panel are shown in the surrounding panels, with two small sub panels for each object. Surrounding panels marked with ‘a’ on the top left side show the overall shape of the spectrum, whereas the panels marked ‘b’ show the  $H\alpha$  region. Each object is labeled in sub panel ‘a’ by SDSS ObjID. Black solid lines indicate the observed spectrum. The two narrow bands that define the flux ratio are shaded gray. Four spectra (numbered from 1 to 4) are chosen as type 1 AGN candidates out of six examples by the demarcation lines. Flux has the unit  $[10^{-17} \text{ ergs}^{-1} \text{ cm}^{-2} \text{ \AA}^{-1}]$ .

and hard X-ray data in the literature (Ueda et al. 2003; Grimes et al. 2004; Hasinger 2004).

A theoretical framework has been claimed to explain such a luminosity dependency of the type 1 AGN fraction. For example, the receding torus model (Lawrence 1991; Falcke et al. 1995; Hill et al. 1996; Simpson 2005) imposes a dust sublimation radius of the torus, which is directly affected by the nuclear luminosity. As the luminous central engine pushes the dust torus away, thereby increasing the torus’ inner radius, the opening angle must increase if the height of the torus is fixed. This naturally explains the observed luminosity dependence of the type 1 AGN fraction.

A precise determination of the type 1 AGN fraction is obviously a key test of a model’s validity. Since the AGN unified model explains the observability of both types of AGN as the consequences of obscuration induced by the geometry and the structure of AGN, the frequency of observed type 1 AGN indirectly constrains the model. For instance, higher type 1 AGN fraction implies wider open-

ing angle, smaller scale height of the torus, and higher nuclear luminosity. Herein we report on our efforts to discover new type 1 AGNs from the SDSS database to provide such a test. We describe how we found a substantial number of new type 1 AGNs and derived a more up-to-date type 1 AGN fraction. The use of the new type 1 AGN fraction places a harsh constraint on the model, and if the scheme used herein proves convincing, it appears that an important modification to the model will be necessary. We release a new catalogue of type 1 AGNs in the nearby universe by means of this paper.

The paper is organized as follows. In Section 2, we describe the selection technique whereby we constructed the new type 1 AGN database, and describe a simulation that quantified the detection efficiency of our method. We demonstrate our conservative scheme for type 1 selection and determine the final type 1 AGN sources, taking the simulation into account. In Section 3, we present the new type 1 AGN fraction and discuss how contamination caused by star formation activities may affect it.

Finally, we discuss our findings and summarize our results in Section 4.

We assume a cosmology with  $h = 0.70$ ,  $\Omega_M = 0.30$ , and  $\Omega_\Lambda = 0.70$  throughout this work.

## 2. TYPE I AGN SAMPLE SELECTION

### 2.1. Selection based on broad H $\alpha$ features

The presence of unidentified type 1 AGNs in the nearby universe has been reported by [Hao et al. \(2005\)](#), [Greene & Ho \(2007\)](#), and [Oh et al. \(2011\)](#) [a.k.a. the OSSY catalogue]. In particular, [Oh et al. \(2011\)](#) found that the presence of broad Balmer emission lines were responsible for a very poor fit by model to both the emission lines and the stellar continuum during their construction of the OSSY catalogue. The authors used a flux ratio near the H $\alpha$  emission line to systematically search for BLR AGNs from the SDSS DR7 galaxy database. In this work, we adopted their basic scheme but developed it further to minimize dubious detections.

The OSSY catalogue provides new spectral line strength measurements for the SDSS DR7 galaxies at  $z < 0.2$  ( $N = 664,187$ ). By combining stellar kinematics measurements (pPXF; [Cappellari & Emsellem 2004](#)) and spectral line fitting algorithm (gandalf; [Sarzi et al. 2006](#)), the authors released improved stellar absorption and emission-line strengths, stellar velocity dispersion, and quality assessing parameters. We start our analysis based on the OSSY catalogue.

To identify the spectra showing broad H $\alpha$  features<sup>6</sup>, we computed a ratio between the mean fluxes at two specific wavelength bands. The first narrow band (6,460 – 6,480 Å) is located near H $\alpha$  but far enough from it to be free from the H $\alpha$  broad emission, mainly to secure the pseudo-continuum for H $\alpha$ . The second (6,523 – 6,543 Å) was placed close to [N II]  $\lambda 6548$  to highlight the broad H $\alpha$  feature. Then, we computed a flux ratio ( $F_{6,533}/F_{6,470}$ ) based on the two bands. Fig. 1 shows the flux ratio versus the signal-to-statistical noise ratio (S/sN) of the chosen specific wavelength bands for all galaxies listed in the OSSY catalogue ( $N = 664,187$ ;  $z < 0.2$ ). We used this diagram to first detect H $\alpha$  BLR candidates. Examples of spectral energy distributions (SEDs) for various values of flux ratios are illustrated in Fig. 1 (surrounding panels), including the overall spectra and the H $\alpha$  regions. As the flux ratio increases, broad H $\alpha$  features become more prominent (see objects marked 1, 3, and 5 in order of decreasing broad H $\alpha$  strength), and the same trend is visible in the higher S/sN regime (objects 2, 4, and 6). It seems clear that the flux ratio is effective for detecting H $\alpha$  BLR candidates.

The  $1\sigma$  demarcation line shown in Fig. 1 was applied to first select type 1 AGN candidates. A total of 17% of galaxies have a flux ratio above this cut. We used gandalf (Gas and Absorption Line Fitting, [Sarzi et al. 2006](#)) to fit the spectrum based on both nebular emissions and stellar absorptions. Our approach for analyzing the SDSS spectra can be summarized as a two-step process. First, we masked nebular emission lines and derived the stellar kinematics by directly matching the observed spectrum with stellar templates by using the pPXF IDL

<sup>6</sup> broad H $\alpha$  is also a signature of supernovae but it is known as rare ([Filippenko 1997](#)).

TABLE 1  
SIMULATION DESIGNS

components	range	increment	No. of bins
S/sN <sup>a</sup>	5 – 55, > 55	5	11
FWHM	50 – 15,000 [km s <sup>-1</sup> ]	50	300
Peak <sup>b</sup>	$0.01 \times A_{H\alpha}^c - 0.09 \times A_{H\alpha}$	0.01	9
	$0.1 \times A_{H\alpha} - 1.0 \times A_{H\alpha}$	0.1	10

NOTE. — In total, 627,000 mock galaxies were generated.

<sup>a</sup> Average S/sN from 4 continuum bands (4,500-4,700, 5,400-5,600, 6,000-6,200, 6,800-7,000 Å)

<sup>b</sup> Peak of the Gaussian

<sup>c</sup> Amplitude of the narrow H $\alpha$

program ([Cappellari & Emsellem 2004](#)). For the “stellar” templates, we used the [Bruzual & Charlot \(2003\)](#) stellar population synthesis models and the MILES stellar library ([Sánchez-Blázquez et al. 2006](#)). After measuring the kinematic broadening, we lifted the masking on the nebular emission lines and simultaneously fit stellar continuum and emission lines by using the Levenberg-Marquardt minimization ([Markwardt 2009](#))<sup>7</sup>.

The standard pPXF and gandalf procedures generally work well on normal galaxies, but a more elaborate consideration is necessary for type 1 AGNs. For example only for type 1 AGNs, we adopted a wide masking width of 12,000 km s<sup>-1</sup> for broad Balmer lines (H $\alpha$ , H $\beta$ , H $\gamma$ , and H $\delta$ ) and 1,200 km s<sup>-1</sup> for narrow emission lines, prior to measuring stellar kinematics. Besides, we imposed an initial guess of a large width, 1,000 km s<sup>-1</sup>, for broad Balmer component fitting. We implemented single broad Gaussian component for each Balmer line. As demonstrated in the middle panels of Fig. 2, these prescriptions were effective for decomposing the broad (H $\alpha$ ) and narrow (H $\alpha$ , [N II]  $\lambda 6548$  and [N II]  $\lambda 6584$ ) components around the broad H $\alpha$  region.

Following these routines, we measured the full-width at half-maximum (FWHM), flux, and amplitude over noise (A/N) ratio<sup>8</sup> of all lines listed in the OSSY catalogue. Our selection criteria for type 1 AGN candidates are as follows.

- $0.00 < z < 0.20$
- FWHM of broad H $\alpha$  > 800 km s<sup>-1</sup>
- A/N of broad H $\alpha$  > 3

The redshift limit is constrained by the wavelength of the red end of the pseudo-continuum around H $\alpha$ . We chose a cut for FWHM(H $\alpha$ ) that was slightly lower than the one more widely used, 1,000 km s<sup>-1</sup> ([Vanden Berk et al. 2006](#); [Schneider et al. 2010](#); [Stern & Laor 2012](#)), hoping to detect even weaker BLR AGNs. As a result, we found 9,671 (8.6%) type 1 AGN candidates among the 111,824 sources that satisfied the  $1\sigma$  cut shown in

<sup>7</sup> pPXF and gandalf use the MPFIT IDL routine provided by Craig Markwardt at <http://cow.physics.wisc.edu/~craigm/idl/>

<sup>8</sup> A/N is defined by the ratio between Gaussian amplitude of the emission line and dispersion of continuum residual of the given spectrum

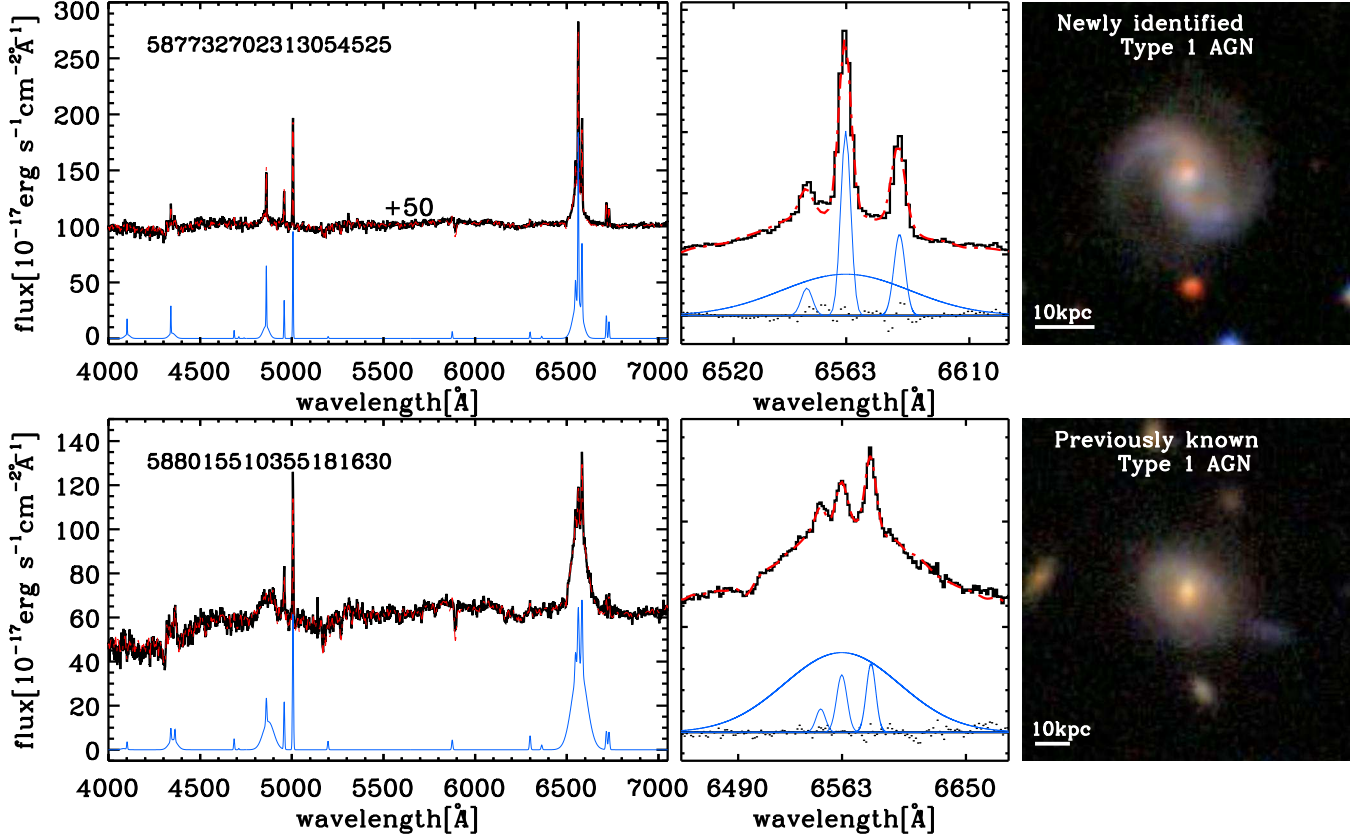


FIG. 2.— Example spectra with SDSS *gri* composite images for a newly identified type 1 AGN (top) and a previously known type 1 AGN from the SDSS pipeline (bottom). Left and middle panels present the overall spectra and their H $\alpha$  regions, respectively. Black lines indicate the observed spectra and red dashed lines include stellar continuum fits. Note that the observed spectrum and the fit in the top panel are shown with offset in the ordinate for clarity (+50 in flux unit). Detected emission lines are shown in the left panels as blue lines. In particular, broad H $\alpha$  components are shown in the middle panels as well as other narrow emission lines. Dots in the middle panels show the residuals.

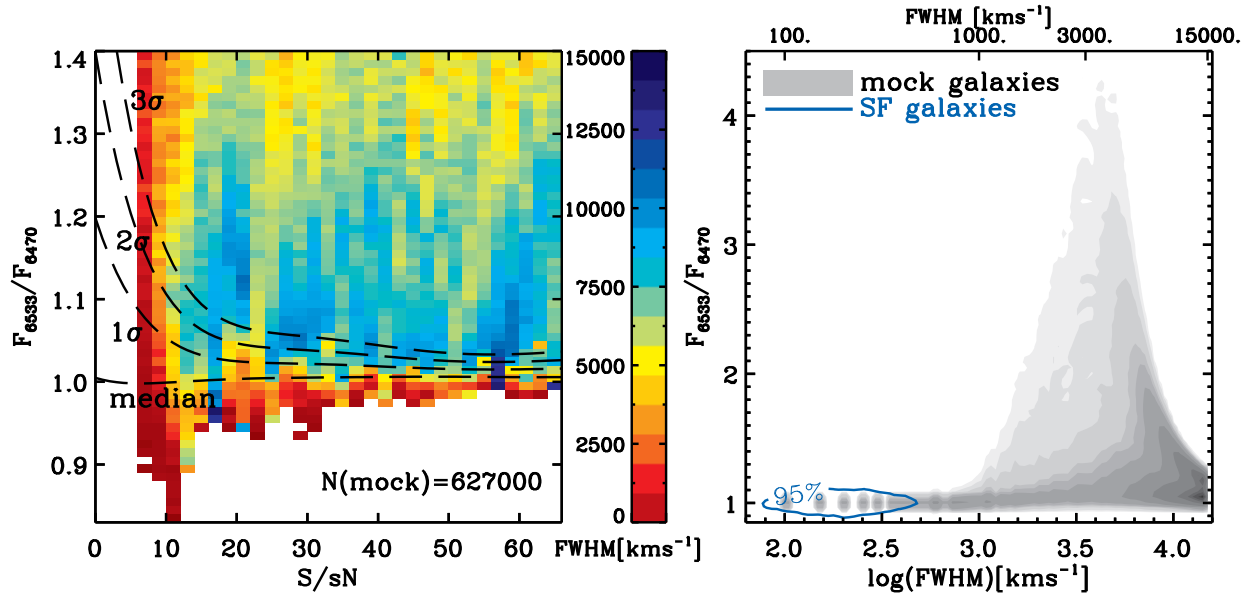


FIG. 3.— Flux ratio distribution with respect to S/N (left) and FWHM (right) for the simulated galaxies. In the left panel, color grids and the corresponding color bar represent the mean FWHM. In the right panel, FWHMs measured for the actively star-forming galaxies are indicated by a blue contour (95%) while the values for the mock spectra are indicated by achromatic filled contours.



TABLE 2  
DETECTION EFFICIENCY FOR EACH  $\sigma$  CUT

$\sigma$ cut <sup>a</sup>	No. of mock galaxies	No. of mock BLRs	completeness <sup>b</sup> (%)	purity (%)
$> 1\sigma$	547,495	543,052	91.49	99.19
$> 2\sigma$	515,939	514,562	86.69	99.73
$> 3\sigma$	487,943	487,512	82.13	99.91

<sup>a</sup>  $\sigma$  cut marked in the central panel of Fig. 1

<sup>b</sup> Number of all mock broad-line AGNs ( $\text{FWHM} > 800 \text{ km s}^{-1}$ ) = 593,560

TABLE 3  
DETECTION EFFICIENCY FOR EACH  $\sigma$  CUT AFTER APPLYING  
AREAL RATIO

$\sigma$ cut	No. of mock galaxies	No. of mock BLRs	completeness (%)	purity (%)
$> 1\sigma$	547,495	484,791	81.68	88.55
$> 2\sigma$	515,939	476,341	80.25	92.33
$> 3\sigma$	487,943	463,475	78.08	94.99

the central panel of Fig. 1. Of these 9,671, only 120 were of  $\text{FWHM}(\text{broad H}\alpha) < 1,000 \text{ km s}^{-1}$ , and hence the choice of the  $800 \text{ km s}^{-1}$  cutoff instead of the more typical  $1,000 \text{ km s}^{-1}$  cutoff influenced the results and conclusions little.

## 2.2. Test of the selection scheme

The selection process described above relies on the specific flux ratio as a proxy to the broad  $\text{H}\alpha$  feature. To verify our approach and measure the detection efficiency, that is, its completeness and purity, we conducted a simulation by generating mock spectra, each with an artificially broadened  $\text{H}\alpha$  component.

The basis of the mock spectra was 110 star-forming galaxies with narrow emission lines: 10 galaxies randomly chosen from each of 11 S/SN bins. To secure a wide range of the amplitude of  $\text{H}\alpha$ , we used the spectra of actively star-forming galaxies of  $\text{SFR} > 0.5 M_{\odot} \text{ yr}^{-1}$  that were identified by the BPT diagnostics diagram (Baldwin et al. 1981) and the Kennicutt law (Kennicutt 1998). We then add broad line components by varying two parameters: Gaussian peak and FWHM, while the center of the Gaussian was fixed at  $6562.8 \text{ \AA}$ . The peak of the Gaussian was fractionally scaled by the amplitude of the narrow  $\text{H}\alpha$  emission line ( $A_{\text{H}\alpha}$ ) in 19 steps, namely, by 1% increments from 1% to 10% and a 10% increment from 10% to 100% in  $A_{\text{H}\alpha}$ . The FWHM of the broad  $\text{H}\alpha$  line was varied between 50 and  $15,000 \text{ km s}^{-1}$  in  $50 \text{ km s}^{-1}$  increments. For reference, the typical range of  $\text{FWHM}(\text{H}\alpha)$  of star-forming galaxies is  $50\text{--}350 \text{ km s}^{-1}$  within the  $2\sigma$  level. In total, 627,000 ( $110 \times 19 \times 300$ ) mock spectra were generated, as summarized in Tab. 1.

Fig. 3 shows flux ratios versus S/SN (left panel) and FWHM (right panel) for the simulated galaxies. In the left panel, the galaxies with low FWHMs (dark red

color) were generally found below the  $1\sigma$  demarcation line and/or had low S/SN. As the color grids show, our  $1\sigma$  cut for recovering BLRs with a large value of FWHM (blue color) seems effective, although it still suffers from contaminants such as objects having noise fluctuations or having a broad  $\text{H}\alpha$  component that is too shallow. We will discuss how these contaminants can be removed in the next section. The right panel of Fig. 3 shows that the highly star-forming galaxies having only narrow emission lines obviously occupy the region of low flux ratio ( $< 1.1$ ) and low FWHM ( $< 500 \text{ km s}^{-1}$ ). Below the FWHM of  $500 \text{ km s}^{-1}$ , the mock galaxies have almost the same flux ratios as narrow emission line galaxies, which means that it is not easy to distinguish between weak type 1 AGNs and star-forming galaxies. On the other hand, the flux ratio of the mock galaxies begins to spread out greatly for  $\text{FWHM} > 1,000 \text{ km s}^{-1}$ .

Detection efficiency (completeness and purity) can be measured by assuming that a BLR AGN has  $\text{FWHM}(\text{broad H}\alpha) > 800 \text{ km s}^{-1}$ . Completeness is defined as the fraction of mock BLR AGNs that are recovered by the procedure (i.e., the  $\sigma$  cut). Purity describes the fraction of BLR AGNs among all galaxies selected by a  $\sigma$  cut. Tab. 2 lists the completeness and purity of BLR AGNs for three choices of  $\sigma$  cut. As higher  $\sigma$  cuts are applied, completeness decreases but purity increases.

Fig. 4 illustrates that the completeness (color key) of our selection method based on the flux ratio of the broad  $\text{H}\alpha$  component is mainly determined by the broad  $\text{H}\alpha$  luminosity (ordinate). We used this completeness distribution to find BLR AGN candidates most economically. We first located the locus of 90% completeness pixels in the diagram. As most of the pixels above this locus satisfy the completeness of 90%, the total completeness when using this 90% cut is in practice much greater than 90%. Roughly 21% (2,052 out of 9,671) of the BLR AGN candidates (by the  $1\sigma$  criterion) were below the 90% demarcation line and thus were missed by this scheme.

## 2.3. Selection based on the broadening around [NII] $\lambda 6584$

We have so far described our selection criterion based on the broad  $\text{H}\alpha$  feature, but there still are complexities. For instance, an extremely shallow broad  $\text{H}\alpha$  component can be technically classified as type 1 AGN as long as it satisfies our conditions on A/N and FWHM but not without doubt. For another example, the credibility of type 1 classification for border-line objects of  $\text{FWHM}(\text{H}\alpha)$  just over  $800 \text{ km s}^{-1}$  can be doubted, because it is difficult to measure the width of such a marginally-broad  $\text{H}\alpha$  component accurately, especially in the midst of strong neighboring narrow emission lines ([N II]  $\lambda 6584$ ).

Having carried out a sophisticated spectral fit for each galaxy spectrum and having determined decomposed broad  $\text{H}\alpha$  properties as a result, we were able to perform a more stringent type 1 classification than carried out so far. We compared the area of the broad  $\text{H}\alpha$  component in question with the noise level of the observed continuum. The area of broad  $\text{H}\alpha$  beyond [N II]  $\lambda 6584$  to longer wavelengths can be used as an alternative measure to FWHM for the broad  $\text{H}\alpha$  feature. Meanwhile, the area defined by the dispersion (measurement uncertainty) of the continuum can be used as a control parameter. Fig. 5

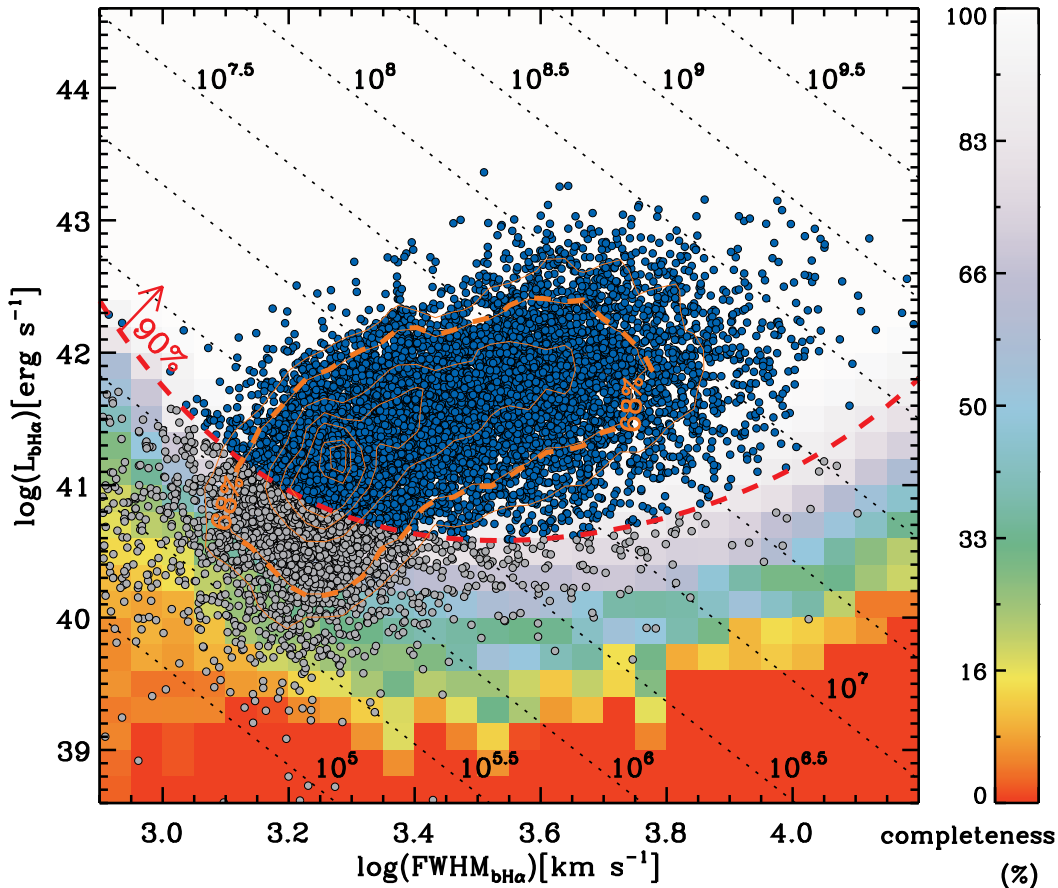


FIG. 4.— Luminosity versus FWHM for broad H $\alpha$  components. Completeness derived from a simulation is indicated by the underlying color grid. Labeled black dotted lines represent  $M_{\text{BH}}$  in units of solar mass following the formulae developed by Greene & Ho (2005). The red dashed curve indicates the demarcation used to choose type 1 AGN sources based on a 90% completeness criterion. The type 1 AGNs ( $N=7,619$ ) selected by the 90% completeness cut are shown as blue-filled dots; excluded points ( $N=2,052$ ) below the curve are shown as gray-filled dots.

provides an illustration with example spectra. The right panels demonstrate how we assess the broad H $\alpha$  component (gray shaded area) that surpasses the level of the continuum dispersion (hatched area), which is marked as  $\sigma_{\text{cont.}}$ . As an additional type 1 AGN criterion, we take the areal ratio of the broad H $\alpha$  over the measurement uncertainty around [N II]  $\lambda 6584$  greater than 2. The application of this new cut removed 3,940 controversial type 1 AGN candidates (Fig. 6). Tab. 3 lists the completeness and purity for each  $\sigma$  cut after implementing areal ratio analysis. As is the case with any other criterion, some of the objects removed in this step are likely to be true type 1 AGN. Yet, we felt it necessary to apply this additional criterion to construct a more robust list of type 1 AGN candidates.

By using the completeness ( $> 90\%$ ) and areal ratio ( $> 2$ ) arguments, we identified 5,553 type 1 AGN candidates. It should be noted that the majority (91.3%) of the type 1 AGN candidates below the 90% completeness cut in Fig. 4 were eliminated through the areal excess test, confirming the effectiveness of the completeness test as a filtering scheme. The flowchart in Fig. 7 shows the number of objects that survived each selection. The broad nature of the H $\alpha$  emission line of type 1 AGNs eluded our initial attempt to measure their strengths.

Perhaps for the same reason, many (38%, 700 out of 1,835) of our newly found type 1 AGNs are not included in the MPA-JHU catalogue<sup>9</sup>. Hereinafter, our analysis is based on these 5,553 sources.

We list the measured properties of these sources as well as the estimated quantities with designation flag in Tab. 9. We used the formula developed by Greene & Ho (2005) that depends on the line width and luminosity of the broad H $\alpha$  to estimate black hole mass. We assumed bolometric luminosity by adopting the relationship developed by Heckman et al. (2004) ( $L_{\text{bol}} \approx 3,500 L[\text{O III}]$ ). The cross-matched flag presents that 456 type 1 AGNs (8.2%) and 2,954 (53.2%) sources in our catalogue are found in Hao et al. (2005) and Greene & Ho (2007), respectively.

#### 2.4. SDSS SpecClass=3 objects

SDSS DR7 provides a spectroscopic classification of galaxies: *SpecClass*. Broad-line AGNs, especially for low redshift, are flagged *SpecClass*=3 when the observed spectrum satisfies the following criteria: the FWHM of

<sup>9</sup> <http://www.mpa-garching.mpg.de/SDSS/DR7/>

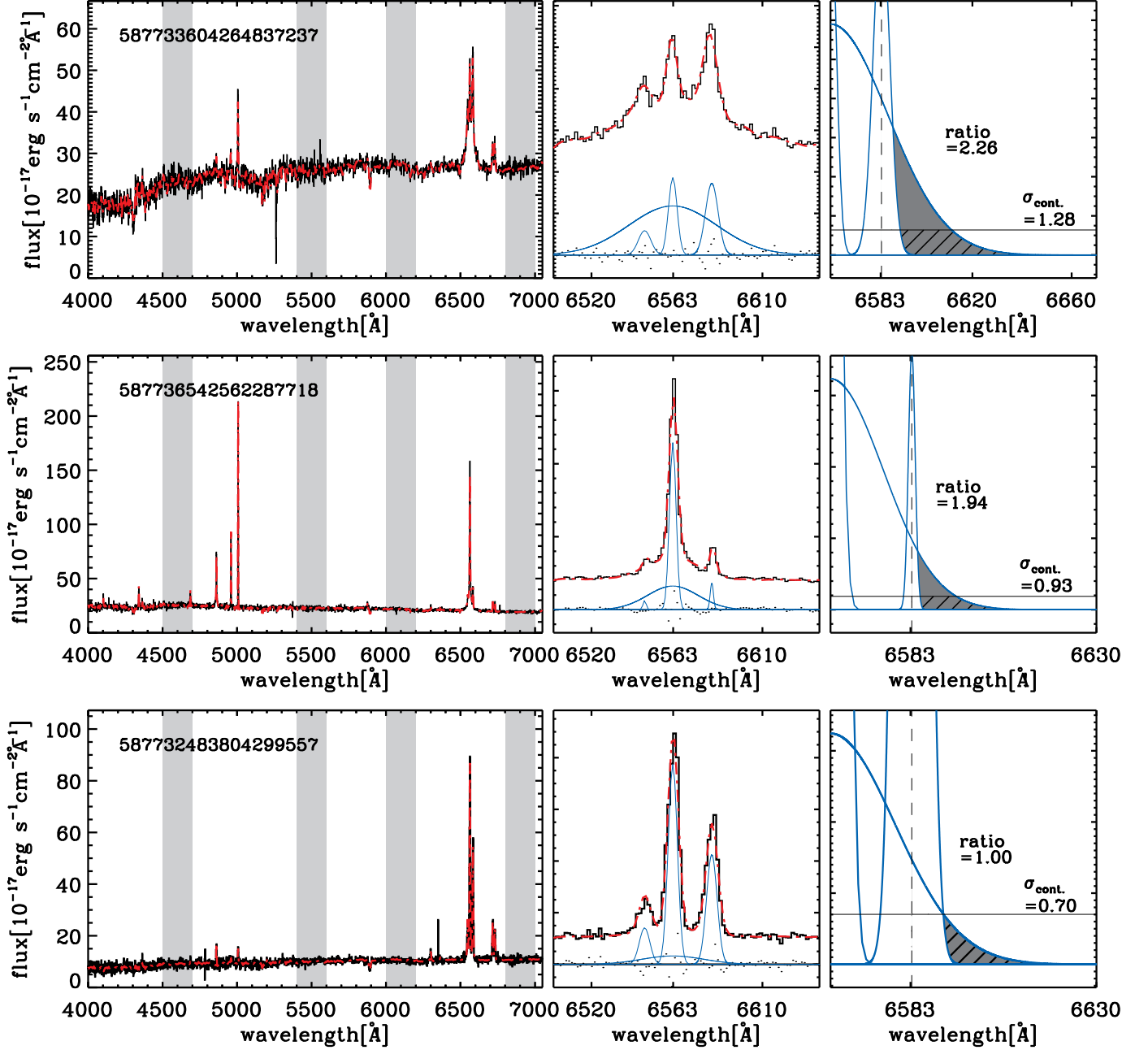


FIG. 5.— Example spectra illustrating the conservative selection criterion based on areal excess. The overall spectra are shown in the left panels, including SDSS ObjID labels at the top left. Black and red lines respectively indicate the observed spectrum and our fit. The regions that are used to measure the averaged continuum dispersion are shaded gray. The middle panels present the detected [N II]  $\lambda 6548$ , narrow H $\alpha$ , broad H $\alpha$ , and [N II]  $\lambda 6584$  emission lines (blue Gaussians), and their residuals (black dots). The right panels enlarge the middle panels, comparing the measured broad H $\alpha$  exceeding [N II]  $\lambda 6584$  (gray shades) to the level of the continuum dispersion (hatched regions with black horizontal lines). The areal excess, that is, the ratio between the gray shaded and hatched regions, is marked, as well as the level of the continuum dispersion,  $\sigma_{\text{cont.}}$ .

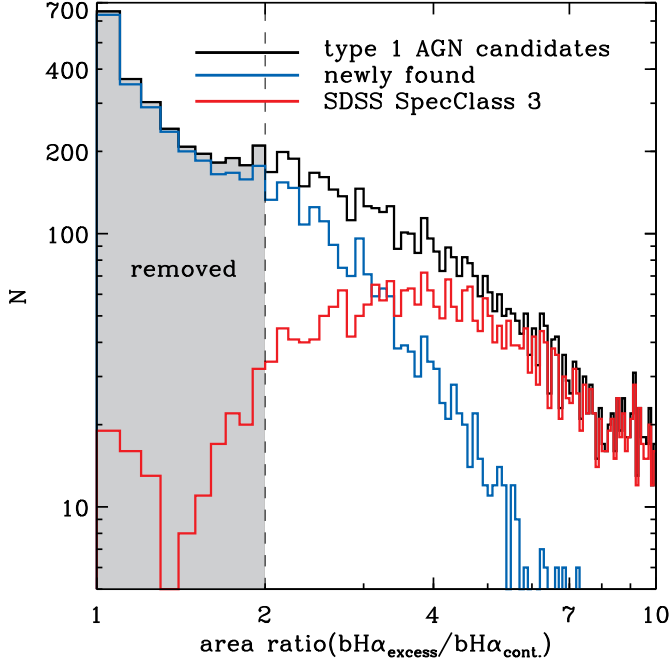


FIG. 6.— Histograms of area ratio for all type 1 AGN candidates (black), newly found type 1 AGNs (blue), and SDSS *SpecClass* 3 sources (red). The new selection criterion, area ratio  $> 2$ , is marked using a vertical dashed line; it removes 2,066 dubious candidates (gray shaded area).

TABLE 4  
DETECTION OF *SpecClass* 3 SOURCES

	N	%
total detected <sup>a</sup>	4125	100.00
after flux ratio cut	4079	98.89
after $A/N_{\text{bH}\alpha}$ & $\text{FWHM}_{\text{bH}\alpha}$ cut	3969	96.22
after completeness cut ( $>90\%$ ) <sup>b</sup>	3921	95.05
after areal excess cut ( $>100\%$ ) <sup>c</sup>	3718 <sup>d</sup>	90.13
missed	407	9.87
Description on the sources that failed flux ratio cut		
only narrow emission lines	19	41.30
absence of spectral information	11	23.91
maybe broad-line AGNs	10	21.74
passing $A/N_{\text{bH}\alpha}$ & $\text{FWHM}_{\text{bH}\alpha}$ cut	(10)	(21.74)
passing completeness cut ( $>90\%$ )	(8)	(17.39)
passing areal excess cut ( $>100\%$ )	(5)	(10.87)
asymmetric bump on $\text{H}\alpha$	3	6.52
non-emission lines	2	4.35
unknown spectrum	1	2.17

<sup>a</sup> Number of all detected *SpecClass* 3 sources chosen from the flux ratio analysis at  $z < 0.2$

<sup>b</sup> see Section 2.3 and Fig. 4

<sup>c</sup> see Section 2.4 and Fig. 5

<sup>d</sup> Number of *SpecClass* 3 sources in the final catalog

any emission line is greater than  $1,000 \text{ km s}^{-1}$ , the height of the emission line is at least 3 times higher than its noise (equivalent to our condition of  $A/N > 3$ ), and the equivalent width of the line is larger than  $10 \text{ \AA}$ . In total, there are 4,125 *SpecClass* 3 sources at  $z < 0.2$ . Galaxies not showing broad-emission lines are classified as *SpecClass* 2.

We tried applying our type 1 AGN selection technique

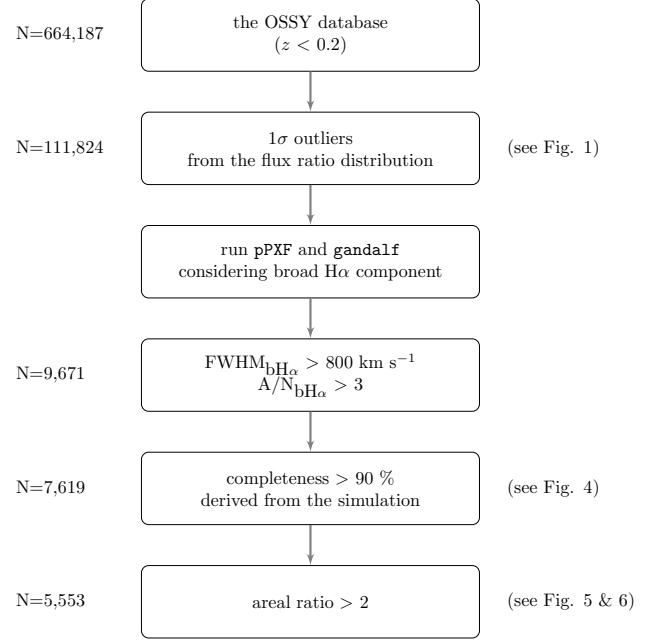


FIG. 7.— Flowchart of the type 1 AGN selection.

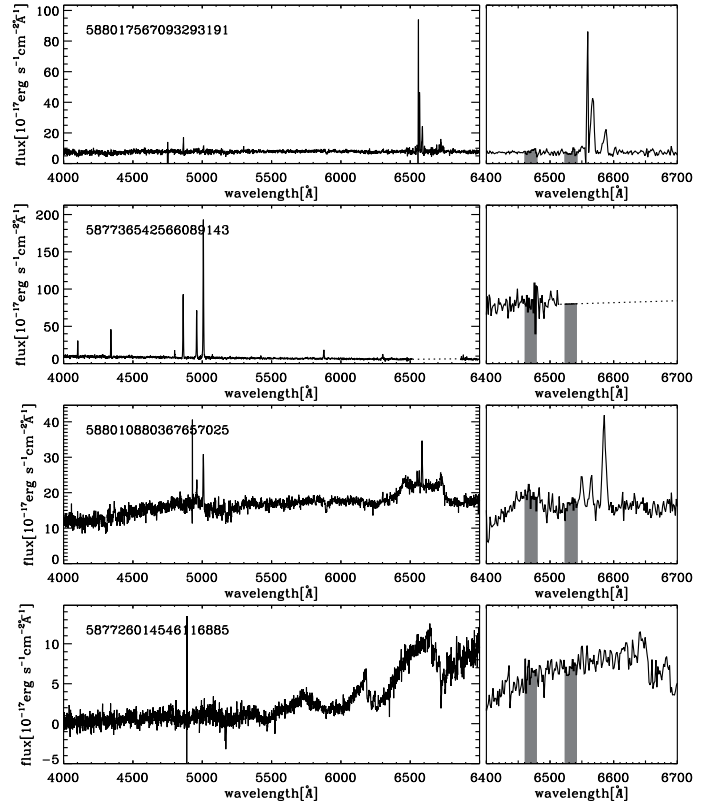


FIG. 8.— Spectra of example *SpecClass* 3 sources below the  $1\sigma$  demarcation line, for a narrow-emission-line galaxy (top), an incomplete spectrum near  $\text{H}\alpha$  (upper), an asymmetric bump feature (lower), and an unknown source (bottom). Right panels show the  $\text{H}\alpha$  region; the two narrow bands that define the flux ratio are shaded gray.



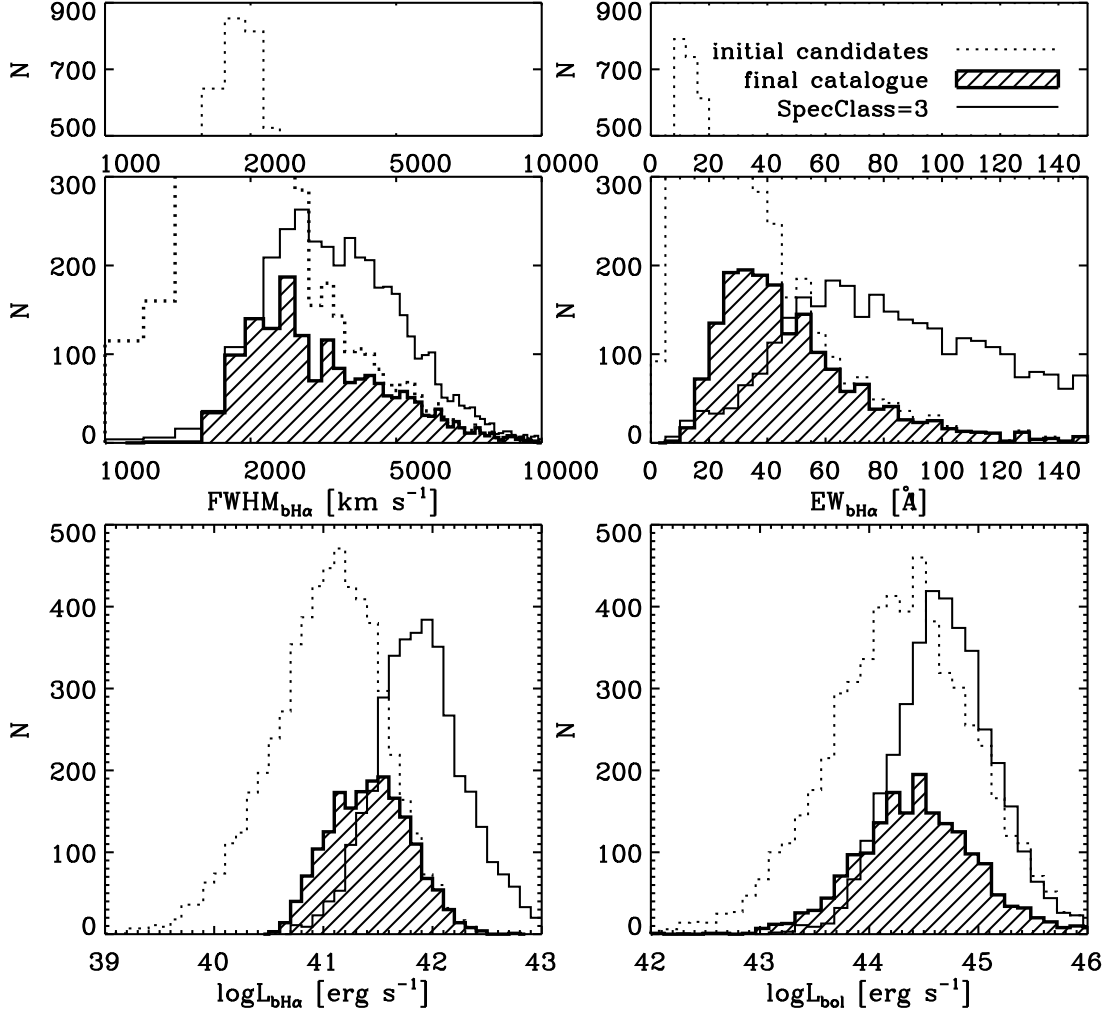


FIG. 9.— Histograms of  $\text{FWHM}_{\text{bH}\alpha}$ ,  $\text{EW}_{\text{bH}\alpha}$ ,  $\log L_{\text{bH}\alpha}$ , and  $\log L_{\text{bol}}$  for newly found type 1 AGNs (hatched area, see Section 2.3), *SpecClass* 3 sources (solid lines), and all type 1 AGN candidates (dotted) that were chosen by measuring the flux ratio  $F_{6,533}/F_{6,470}$  (see Section 2.1).

on the *SpecClass* 3 objects; Tab. 4 summarizes our results. To begin with, 4,079 out of 4,125 objects passed our flux ratio test. Then, 3,969 of the remaining objects passed the A/N and FWHM cuts, and 3,718 (90.13%) passed the final areal excess test. Overall, our technique rejected roughly 10% (407 out of 4,125) of the objects.

We describe further why some of the SDSS *SpecClass* 3 objects are not classified the same way in our analysis based on a subsample of the objects that were excluded by the first and most important step of our scheme, the flux ratio test. By performing additional spectral fitting to the 46 (4,125 – 4,079) sources that failed to pass the  $1\sigma$  cut from the flux ratio distribution, we found that 11 objects have empty spectral bins around H $\alpha$  (see Fig. 8, second row). Hence, our technique could not be feasibly applied to these objects. In the majority of the rest 35 missed objects (25 out of 35), we could not detect any notable broad features; the top row of Fig. 8 shows such an example. Of the 46 missed objects, 10 passed our A/N and FWHM tests of H $\alpha$  and thus possibly have broad H $\alpha$  components. However, 2 of the 10 failed to pass our 90% completeness test and a further 3 failed to pass the areal excess cut test; overall, only 5 out of 10 passed all of

our test. Similarly, our technique could not be effectively applied on 3 objects because their spectrum around H $\alpha$  appeared asymmetric and uneven (see Fig. 8, third row). Finally, one object had a spectrum that did not appear to be an ordinary galaxy spectrum, making it difficult to fit this spectrum with a combination of template stellar (population) spectra (Fig. 8, bottom row). Tab. 4 summarizes the statistics of the *SpecClass* 3 sources.

Fig. 9 presents the distributions of the widths and luminosities of the broad H $\alpha$  component for the newly found type 1 AGNs (hatched area) and *SpecClass* 3 sources (solid line). The new sources show broad H $\alpha$  widths ( $\text{FWHM}_{\text{bH}\alpha}$ ) comparable to those of the *SpecClass* 3 sources. On the other hand, the characteristic luminosity (i.e., the peak of the distribution) and the equivalent width of their broad H $\alpha$  components seems lower, by a factor of a few, than that of the *SpecClass* 3 objects, where bolometric luminosity is derived from  $L_{\text{bol}} \approx 3,500 L[\text{O III}]$  (Heckman et al. 2004). Hence, the newly found objects are mostly low-luminosity type 1 AGNs.

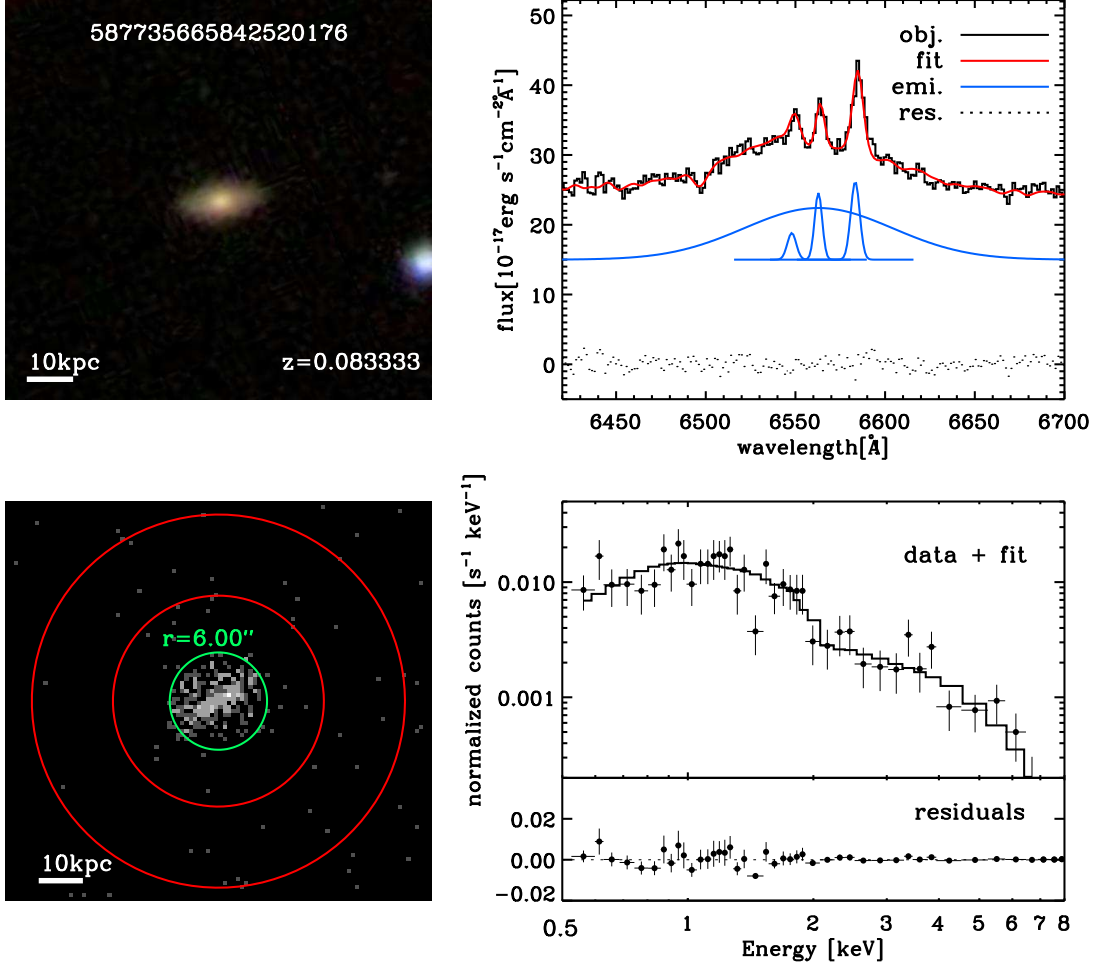


FIG. 10.— Example *Chandra* X-ray data (obsid = 2779) and model fit, with the SDSS *gri* composite image and our spectral fit. Top left: SDSS *gri* composite image. Top right: observed spectrum (black), our fit (red), detected emission lines (blue) including broad components, and residuals (black dots). Note that the detected emission lines are arbitrarily shifted towards the y-axis for clarity. Bottom left: *Chandra* X-ray image, including source (green circle) and background (red annulus) markers. Bottom right: extracted data with the applied fit versus energy (top); residuals (bottom).

TABLE 5  
X-RAY FITTING STRATEGIES

count	fitting prescription $N_{\text{H}}$ [cm <sup>-2</sup> ] <sup>a</sup>	$\Gamma$ <sup>b</sup>
$\leq 50$	$10^{20}$ fixed <sup>c</sup>	1.9 fixed
$50 < \text{count} < 100$	$10^{20-24}$ variable	1.9 fixed
$\geq 100$	$10^{20-24}$ variable	1.5-2.5 variable

<sup>a</sup> Hydrogen column density

<sup>b</sup> X-ray photon index

<sup>c</sup> The value is fixed during spectral fit

### 2.5. *Chandra* X-ray data

A strong X-ray flux is one of the prominent and characteristic features of an AGN, because an accretion disk surrounding a central black hole produces X-ray emissions (Fabian et al. 1989). Relativistic electrons from the corona in the vicinity of the accretion disk scatter the lower-energy photons radiated from the accretion disk. As a result, the lower-energy photons receive more energy from the electrons (inverse Compton scattering), and this

process produces an X-ray spectrum with a power-law continuum (Haardt & Maraschi 1991; Zdziarski et al. 2000; Kawaguchi et al. 2001). Thus, strong X-ray emission is considered to be evidence of the presence of an AGN.

We utilized the *Chandra* Source Catalog (Evans et al. 2010) and archival *Chandra* X-ray data to identify X-ray sources among our AGN candidates. We downloaded the matched X-ray images from the *Chandra* archive and applied a 10 arcmin matching radius. Through visual inspection of X-ray images of their optical counterpart, we found 84 X-ray sources among the type 1 AGN samples. The X-ray recovery rate (1.5%) was very low mainly because X-ray sky coverage was much narrower (by a factor of 25) and depth was much shallower compared to the SDSS optical survey.

We used the *Chandra* Interactive Analysis of Observations (CIAO) v4.6<sup>10</sup> and XSPEC v12.8.1g software<sup>11</sup> (Arnaud 1996) to reprocess the *Chandra* X-ray data and

<sup>10</sup> <http://cxc.harvard.edu/ciao/>

<sup>11</sup> <http://heasarc.nasa.gov/xanadu/xspec/>

TABLE 6  
X-RAY PROPERTIES OF TYPE 1 AGNs

SDSS ObjID	obsid <sup>a</sup>	RA <sup>b</sup> [deg]	DEC <sup>b</sup> [deg]	z	counts	count rate [count s <sup>-1</sup> ]	log $L_{2-10\text{keV}}$ <sup>c</sup> [erg s <sup>-1</sup> ]	note <sup>d</sup>
587738372207476987	3784	121.94200	21.23800	0.14215	230	0.012	42.76 <sup>+0.08</sup> <sub>-0.06</sub>	3
587735665842520176	2779	211.12100	54.39800	0.08333	318	0.022	42.56 <sup>+0.06</sup> <sub>-0.06</sub>	N
587725551741370430	2033	143.96500	61.35300	0.03933	332	0.007	41.36 <sup>+0.04</sup> <sub>-0.04</sub>	N
587741602568798223	4695	184.53300	28.17600	0.17773	449	0.045	43.30 <sup>+0.05</sup> <sub>-0.05</sub>	N
587725551197683729	7146	124.83400	50.00700	0.13969	171	0.022	42.77 <sup>+0.09</sup> <sub>-0.09</sub>	3
587729774220738649	9557	210.21899	-1.75300	0.14878	1662	0.034	42.92 <sup>+0.02</sup> <sub>-0.02</sub>	N
587731521745518723	827	140.28600	45.64900	0.17448	4617	0.246	44.16 <sup>+0.02</sup> <sub>-0.01</sub>	3p
587729653959033079	887	249.20200	41.02900	0.04737	287	0.004	41.28 <sup>+0.05</sup> <sub>-0.07</sub>	N
587735696453009410	1623	220.76199	52.02700	0.14121	2851	0.190	43.86 <sup>+0.02</sup> <sub>-0.02</sub>	3p

NOTE. — (This table is available in its entirety in a machine-readable form in the online journal. A portion is shown here for guidance regarding its form and content.)

<sup>a</sup> *Chandra* observation ID

<sup>b</sup> *Chandra* J2000 coordinates

<sup>c</sup> Errors are estimated with the 90% confidence region for a single interesting parameter.

<sup>d</sup> Discriminating flag for the newly identified type 1 AGNs ('N') and the previously confirmed sources ('3'). Notation 'p' indicates the photon pileup fraction greater than 10%.

perform a spectral fitting analysis. We manually aligned the coordinates to the actual image center and determined the source aperture,  $A_s$ , (green circle in Fig. 10) by reading off the value from the pipeline relation between the off-axis angle and the aperture size suggested. The applied aperture ( $\overline{A_s} = 11.2\text{kpc}$ ,  $\sigma_{A_s} = 10.0\text{kpc}$ ) generally covers the whole galaxy ( $\text{PetroR90}_r = 12.2\text{kpc}$ ,  $\sigma_{\text{PetroR90}_r} = 6.2\text{kpc}$ , where  $\text{PetroR90}_r$  is the radius containing 90% of Petrosian flux in  $r$ -band) while some nearby sources represent central part. Then, we used `specextract` with `psfcorr` in CIAO which incorporates an energy-dependent aperture correction (average enclosed energy fraction is 85%). We set the sky background to an annulus of a 10 arcsec width (or 13 arcsec away from the center when  $A_s < 5$  arcsec (or  $A_s \geq 5$  arcsec)). We grouped all spectra with a minimum of 1 count per bin and applied the Cash statistic (Cash 1979) with the assumption of the Poisson distribution to achieve spectral fitting for low-count spectra. XSPEC command `statistic cstat` applies the W statistic (Wachter et al. 1979) for the unmodeled background spectra. We performed spectral fitting by applying the photoelectric absorption model and power law (`phabs*power` in XSPEC) in the rest-frame energy range between 0.5 and 8.0 keV. Tab. 5 gives a detailed prescription for the spectral fitting procedure, Tab. 6 describes the X-ray properties measured on these sources, and Fig. 10 shows an example result.

Apart from AGNs, there are other luminous X-ray sources such as high-mass X-ray binaries, young supernova remnants, and hot ionized interstellar medium resulting from star formation activity. In-depth studies have proposed a correlation between the global star formation rate (SFR) and X-ray luminosity across the whole galaxy (Grimm et al. 2003; Ranalli et al. 2003; Grimes et al. 2005; Lehmer et al. 2010; Mineo et al. 2011). In particular, Ranalli et al. (2003) showed the following linear relationship between 2–10 keV X-ray luminosity and SFR based on 23 *bona fide* star-forming galaxies.

$$\text{SFR}[M_{\odot}\text{yr}^{-1}] = 2.0 \times 10^{-40} L_{2-10\text{keV}} \quad (1)$$

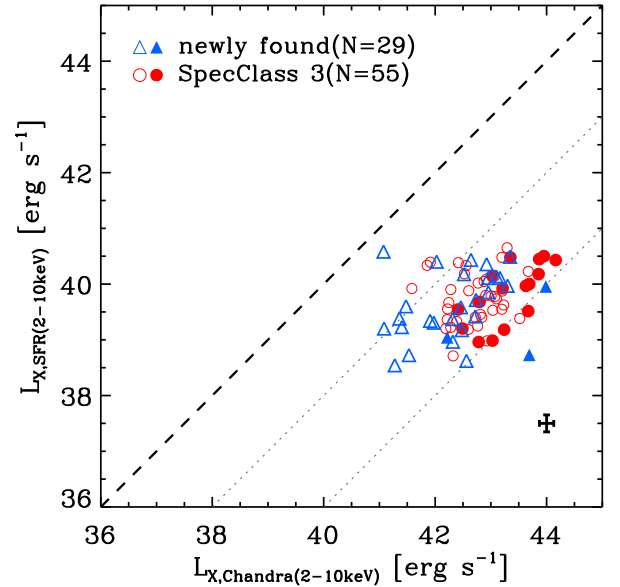


FIG. 11.— Comparison of 2–10 keV X-ray luminosity between measurements ( $L_{X, \text{Chandra}}$ ) and SFR-inferred estimates ( $L_{X, \text{SFR}}$ ). Newly found type 1 AGNs and *SpecClass* 3 sources are shown as blue triangles and red circles, respectively. The objects having more than 10% of photon pileup fraction are shown with filled blue triangles (3 out of 29) and filled red circles (16 out of 55). The black cross at the bottom right indicates the mean error for all sources. The central dashed line is a one-to-one fiducial line. Two additional dotted lines are shown to facilitate comparison.

We can also estimate SFR by using  $\text{H}\alpha$  narrow emission line strength following Kennicutt (1998), and then derive the 2–10 keV X-ray luminosity ( $L_{X, \text{SFR}}$ ) by using this equation. Fig. 11 compares  $L_{X, \text{SFR}}$  and  $L_{X, \text{Chandra}}$  for the 84 type 1 AGNs for both the newly found type 1 AGNs (blue triangles) and the SDSS *SpecClass* 3 sources (red circles). Fluxes (horizontal axis) are calculated from the best-fit model. The *Chandra* X-ray luminosities

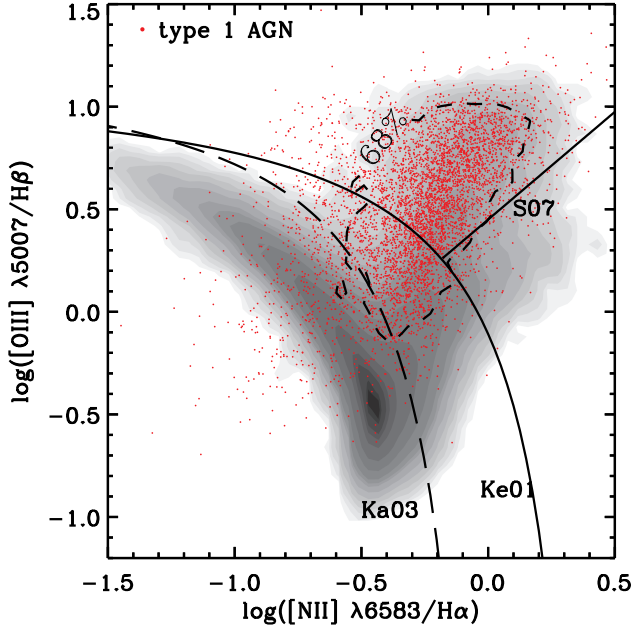


FIG. 12.— BPT diagnostic diagram for type 1 AGNs (red dots). Emission-line galaxies with  $A/N$  cut ( $> 3$  for  $[\text{N II}] \lambda 6584$ ,  $\text{H}\alpha$ ,  $[\text{O III}] \lambda 5007$ , and  $\text{H}\beta$ .  $N \approx 180,000$ ) chosen from the entire OSSY catalogue are shown with filled contours. Dashed contour indicates 68% distribution for type 1 AGNs.

( $L_{\text{X,Chandra}}$ ) are 2–4 orders of magnitude higher than the SFR-inferred estimates ( $L_{\text{X,SFR}}$ ), robustly confirming that our new candidates are indeed AGNs.

Regarding photon pileup fraction of our *Chandra* sources, we performed an investigation based on count rate. We explicitly marked the sources with filled symbols in Fig. 11 if the object has count rate per second greater than 0.07 which corresponds to a photon pileup fraction greater than 10%. Loss of photon count underestimates X-ray luminosity (horizontal axis of Fig. 11). There are 19 sources (out of 84) that show photon pileup fraction greater than 10%, and their intrinsic X-ray luminosity would be found further away from the fiducial line in Fig. 11. Hence, photon pileup does not change our conclusion that 84 X-ray sources are powered by AGN.

### 2.6. Type 2 AGN selection

Based on the OSSY catalogue ( $N = 664,187$ ,  $z < 0.2$ ), we chose the type 2 AGNs from the Seyfert region of the BPT AGN diagnostic diagram (Baldwin et al. 1981). We used the theoretical maximum starburst model of Kewley et al. (2001) and the empirical star formation curve derived by Kauffmann et al. (2003) as demarcation lines. For the discrimination of low-ionization nuclear emission-line regions from Seyfert AGNs, we used the empirical demarcation of Schawinski et al. (2007). In addition, we used the criterion  $A/N > 3$  as a threshold to secure statistical significance for the narrow emission lines ( $[\text{N II}] \lambda 6584$ , narrow  $\text{H}\alpha$ ,  $[\text{O III}] \lambda 5007$ , and  $\text{H}\beta$ ) used in the diagnostics (Fig. 12). As a result, we selected 10,108 type 2 AGNs at  $z < 0.2$ . The BPT AGN diagnostics diagram for type 1 AGNs is presented in Fig. 12. Overall distributions of the type 1 and type 2 AGNs versus redshift are shown in Fig. 13.

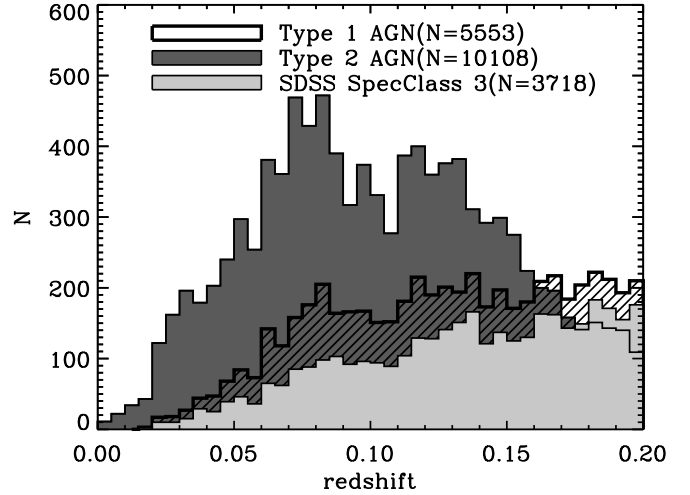


FIG. 13.— Histograms of type 1 (thick line) and type 2 AGNs (dark gray area) with respect to redshift; the type 1 AGN samples include 3,718 *SpecClass* 3 sources (light gray area). The hatched area represents the newly found type 1 AGNs.

## 3. TYPE 1 AGN FRACTION

By combining both types of AGN, we obtained the type 1 AGN fraction, namely the number ratio between type 1 AGNs and all AGNs, as a function of  $[\text{O III}] \lambda 5007$  luminosity and bolometric luminosity (Fig. 14, Tab. 7). The type 1 AGN fraction observed (blue crosses in Fig. 14) increases with  $L[\text{O III}]$ . This trend is generally consistent with the prediction based on the receding torus model, but does not match it closely: the luminosity dependency of the type 1 AGN fraction is steeper at low  $L[\text{O III}]$  and becomes flattened at high  $L[\text{O III}]$  bins compared to the “standard” model (black solid curve) that was derived from the fit to the type 1 AGN fraction from the SDSS DR4. Simpson (2005) noted that the degree of match (to the old data) improved when the height of the torus was modified to depend on the nucleus luminosity in the form of  $h \propto L^\xi$  (black dashed line) with  $\xi > 0$ . In the “modified” scheme, the sublimation radius of a dust torus gradually increases with the nuclear luminosity  $L[\text{O III}]$ . In this framework, the type 1 fraction is described as  $f_1 \equiv N_1/N_{\text{tot}} = 1 - [1 + 3(L/L_0)^{1-2\xi}]^{-\alpha}$ , where  $\alpha$  is set to be 0.5. In Simpson’s formulation, the best fit is to reproduce  $f_1 = \alpha$  for a combination of ( $L_0$ ,  $\xi$ ); this yielded  $L_0 = 42.37 \text{ erg s}^{-1}$  and  $\xi = 0.23$  for the reduced  $\chi^2_\nu$  of 1.5. The modified model does not seem consistent with our new data, however. We will revisit this issue in the next section.

Challenges remain in accurately defining the true type 1 fraction. For example, as discussed in previous sections, our type 1 selection scheme probably excludes some true type 1 AGNs as a consequence of finding robust type 1 AGNs. In this sense, our type 1 fraction may be considered a lower limit. On the other hand, the number of type 2 AGNs is also uncertain. Most notably, we have selected type 2 AGNs only from the Seyfert category based on the BPT diagnostics (see Section 2.5). But, it is likely that there are some type 2 AGNs hidden in the composite and even the star-forming categories. We now attempt to discover them and to include them in our estimation of the type 1 fraction.



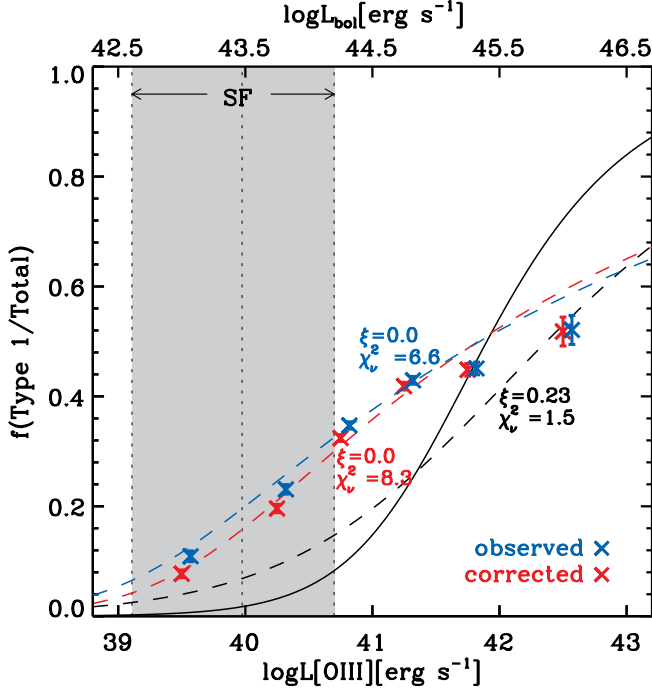


FIG. 14.— Type 1 AGN fraction with respect to  $[\text{O III}] \lambda 5007$  luminosity. Bolometric luminosity is shown on the top axis, following  $L_{\text{bol}} \approx 3500 L[\text{O III}]$  (Heckman et al. 2004). Blue crosses and error bars show the observed type 1 AGN fraction. Red crosses and error bars show the type 1 AGN fraction estimated by statistically correcting for the dilution of observed type 2 AGNs due to star formation. The gray band shows the range of  $L[\text{O III}]$  of star-forming galaxies sampled so that their stellar mass distribution matched that of type 2 AGNs. Vertical dotted lines denote the median and 95% distributions. The black solid curve represents the standard receding torus model proposed by Simpson (2005). The black dashed curve indicates the modified receding torus model which allows to vary the height of the torus according to the nucleus luminosity with  $\xi=0.23$ . The best fits for the observed (blue) and dilution-corrected (red) data points are shown as blue and red dashed lines, including  $\xi$ . Note that the error bars are barely visible due to their small magnitudes, except for the highest  $L[\text{O III}]$  bin.

### 3.1. Dilution correction on optical lines of type 2 AGNs

Our analysis starts by setting up two groups of galaxies at  $z < 0.2$ : “emission-line” galaxies ( $N \sim 200,000$ ) and star-forming galaxies ( $N \sim 150,000$ ) from the OSSY catalog. Emission-line galaxies obviously include star-forming galaxies. We use only the galaxies with  $A/N > 3$  for all four lines of  $[\text{N II}] \lambda 6584$ ,  $\text{H}\alpha$ ,  $[\text{O III}] \lambda 5007$  and  $\text{H}\beta$  that are used in the BPT diagnostics. Assuming that the amount of emission fluxes that come from star formation depends mainly on its stellar mass, we matched every emission-line galaxy with a BPT-selected star-forming galaxy with the same stellar mass within a 5% tolerance level, based on the stellar mass listed in the MPA-JHU catalog. The size (number of galaxies) of the pool of galaxies with (roughly) the same stellar mass for each galaxy used in our test, which depends strongly on stellar mass itself and the shape of stellar mass function, is roughly 3,400 (median).

This match between an emission-line galaxy and a

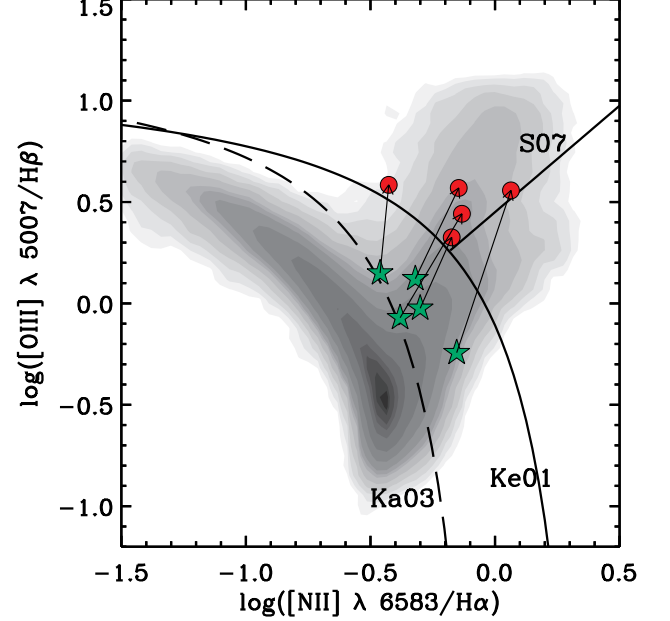


FIG. 15.— BPT diagnostic diagram with 5 randomly chosen examples. All emission line galaxies at  $z < 0.2$  having  $A/N > 3$  on  $[\text{N II}] \lambda 6584$ ,  $\text{H}\alpha$ ,  $[\text{O III}] \lambda 5007$ , and  $\text{H}\beta$  are represented by filled contours. Composite objects (green stars) are shown that were moved to the Seyfert region (red circles) by the procedure described in the main text.

star-forming galaxy solely based on stellar mass is invalid when it is performed on a single galaxy, because even for the same stellar mass of galaxies, star formation rates can vary widely, resulting in different emission line strengths. However, when this exercise is applied to numerous galaxies, and as long as there is no serious bias of star formation rate in any sense in the sample, the method can be considered largely valid. We subtracted the emission line strength of the star-forming galaxy from the emission-line galaxies, for all four narrow lines, and we adjusted the position of each galaxy in the BPT diagram based on its residual line strengths. Note that we have done this on all emission-line galaxies including star-forming galaxies themselves.

After removing star-forming components in the emission lines, 550 star-forming galaxies were moved to the Seyfert region of the BPT diagram, as shown in Fig. 15. This does not mean that there is a significant AGN component in all of these star-forming galaxies. Instead, this arose from the randomness of the choice of our star-forming galaxy as a subtraction. Likewise, after this subtraction, many star-forming galaxies ended up having negative emission lines, again as a result of the randomness of the process. Hence, we decided to ignore this result.

The light emitted by “composite” galaxies is believed to include the contributions of both star-forming and AGN activities, and thus this exercise makes more sense when applied to these galaxies; applying this process moved 1,374 composite galaxies from the composite region to the Seyfert region (Fig. 15). Again, some of these moves may be a result of applying an unfair star-formation correction due to the randomness of our pro-

TABLE 7  
THE OBSERVED TYPE 1 AGN FRACTION WITH THE DILUTION CORRECTED DATA

$\log L_{\text{bol}}$ [erg s <sup>-1</sup> ]	$\log L[\text{O III}]$ [erg s <sup>-1</sup> ]	f(type 1 AGNs/total) [%]	f(type 1 AGNs/total) <sup>a</sup> <sub>corrected</sub> [%]	N <sub>type 1 AGN</sub>	N <sub>type 2 AGN</sub>	N <sub>corrected</sub> <sup>b</sup> <sub>type 2 AGN</sub>
42.5 – 43.5	39.0 – 40.0	10.90 ± 0.98	7.72 ± 0.71	110	899	416
43.5 – 44.0	40.0 – 40.5	23.08 ± 0.86	19.60 ± 0.75	553	1843	426
44.0 – 44.5	40.5 – 41.0	34.71 ± 0.67	32.42 ± 0.63	1778	3344	363
44.5 – 45.0	41.0 – 41.5	42.92 ± 0.71	41.86 ± 0.70	2084	2772	123
45.0 – 45.5	41.5 – 42.0	45.08 ± 1.16	44.83 ± 1.15	833	1015	10
45.5 – 46.5	42.0 – 43.0	52.09 ± 2.64	51.80 ± 2.63	187	172	2

<sup>a</sup> Dilution corrected type 1 AGN fraction which is described in the Section 3.1

<sup>b</sup> Number of dilution corrected type 2 AGNs.

cess. This yielded a total of 11,482 (10,108 + 1,374) type 2 AGNs. The new type 1 AGN fraction that resulted from this “dilution” correction is shown in Fig. 14, using red crosses and error bars.

The analysis is not complete, because we randomly chose and considered emission line fluxes of star-forming galaxies that only matched the stellar mass. Nevertheless, we can expect a general trend of star-formation-caused dilution of AGN relying on the large-number statistics. The finding of new type 2 AGNs reduces the type 1 AGN fraction, especially in low-luminosity bins, but it seems clear that the current versions of the receding torus model still do not reproduce the “corrected” type 1 fraction.

We now attempt to find fits to our new data by adopting the general form of Simpson’s formula, but treating  $\alpha$ ,  $L_0$  and  $\xi$  as free parameters. We assume  $\xi \geq 0$  because this condition is essential to the receding torus model. The best fits were found with  $\alpha = 0.12$ ,  $L_0 = 39.7$ ,  $\xi = 0.0$ , and  $\chi^2_\nu = 6.6$  for the “observed” data set and with  $\alpha = 0.14$ ,  $L_0 = 40.1$ ,  $\xi = 0.0$ , and  $\chi^2_\nu = 8.3$  for the “corrected” data set. The new fits are shown as blue and red dashed lines in Fig 14. However, these two best fits have goodness of fit values ( $\chi^2_\nu$ ) that are too large to accept, and are much larger than that found by Simpson for the old data (1.5). In our test,  $\chi^2_\nu$  became smaller with decreasing  $\xi$ . Our best fit was found with  $\xi = 0$ . This implies that the introduction of “modification” (i.e., a positive luminosity dependence of torus height) seems unsupported. In fact better fits were achieved with a negative value of  $\xi$ ; but the improvement was marginal. The new data call for a new and improved explanation.

#### 4. SUMMARY & DISCUSSION

We have presented a process for finding new type 1 AGNs at  $z < 0.2$  from the SDSS DR7 database. Adding the new type 1 AGNs (1,835) to those in the previous database (3,718) that also pass our selection criteria increases the size of the type 1 AGN database by 49%. Based on the new database, we calculated a new type 1 AGN fraction, which provides a useful test and constraint to the theoretical understanding of the AGN geometry. We are pleased to release the spectroscopic properties of all type 1 AGNs, given in Tab. 9, and our calculations of the type 1 AGN fraction based on  $L[\text{O III}]$  and the Eddington ratio, given in Tab. 7 and Tab. 8, respectively. It is our wish that these tables will be useful to the community.

We showed that our updated type 1 AGN fraction is not consistent with the current versions of the receding

torus model. The type 1 AGN fraction in the low and intermediate luminosity bins is significantly higher than that given by previous data and models, even after correcting for dilution effects on emission lines caused by star formation. We found fits to the new data based on the basic formulation of the receding torus model, allowing large differences in the parameterization process, but these fits were poor. Thus, the new fits do not seem to support the hypothesis that torus depth depends on AGN luminosity.

We present the type 1 AGN fraction with respect to black hole mass and bolometric luminosity in Fig. 16. We used the same bolometric correction factor, 3,500, to both types of AGN following Heckman et al. (2004). The bolometric correction factor may be uncertain; but if there is no serious type bias on it, our results and conclusions in the relative sense would hold good. The color grids in the main panel of the Fig. 16 show that there is a favored area in which the type 1 AGN fraction is high, for higher black hole mass and bolometric luminosity. This ridge-shaped distribution suggests that neither black hole mass nor bolometric luminosity solely determines the type 1 AGN fraction. Rather, the type 1 AGN fraction suggests a rising and falling distribution versus the Eddington ratio (see small panels in Fig. 16). If this trend is real, it may imply that the opening angle initially widens with increasing Eddington ratio until the Eddington limit, and diminishes thereafter, which is not expected based on the basic scheme of the receding torus model. Admittedly, however, the goodness of fits ( $\chi^2_\nu$ ) is too poor to draw any definite conclusion yet.

The structure of the dust torus is a key ingredient of understanding the AGN unification theory. However, its physical properties and the detailed geometry are still under debate. Krolik & Begelman (1988) proposed that the dust and gas in the torus take the form of optically thick, self-gravitating clouds. A clumpy and obscuring torus requires a small volume filling factor and a sufficient scale height compared to the mean free path of clouds (Hönig & Beckert 2007). In the framework of a clumpy torus model, the largest clouds in the torus become gravitationally unbound from the high-luminosity AGN when the AGN accretes and radiates close to the Eddington limit. Based on this argument, the relationship  $H/R \sim L^{-1/4}$  was derived where  $H$  is the scale height,  $R$  is inner radius of the torus, and  $L$  is the AGN luminosity. Note here that the scale height inversely correlates with the AGN luminosity unlike in the receding torus model. Elitzur & Shlosman (2006) described the origin of dust clouds as an outflow wind released from an

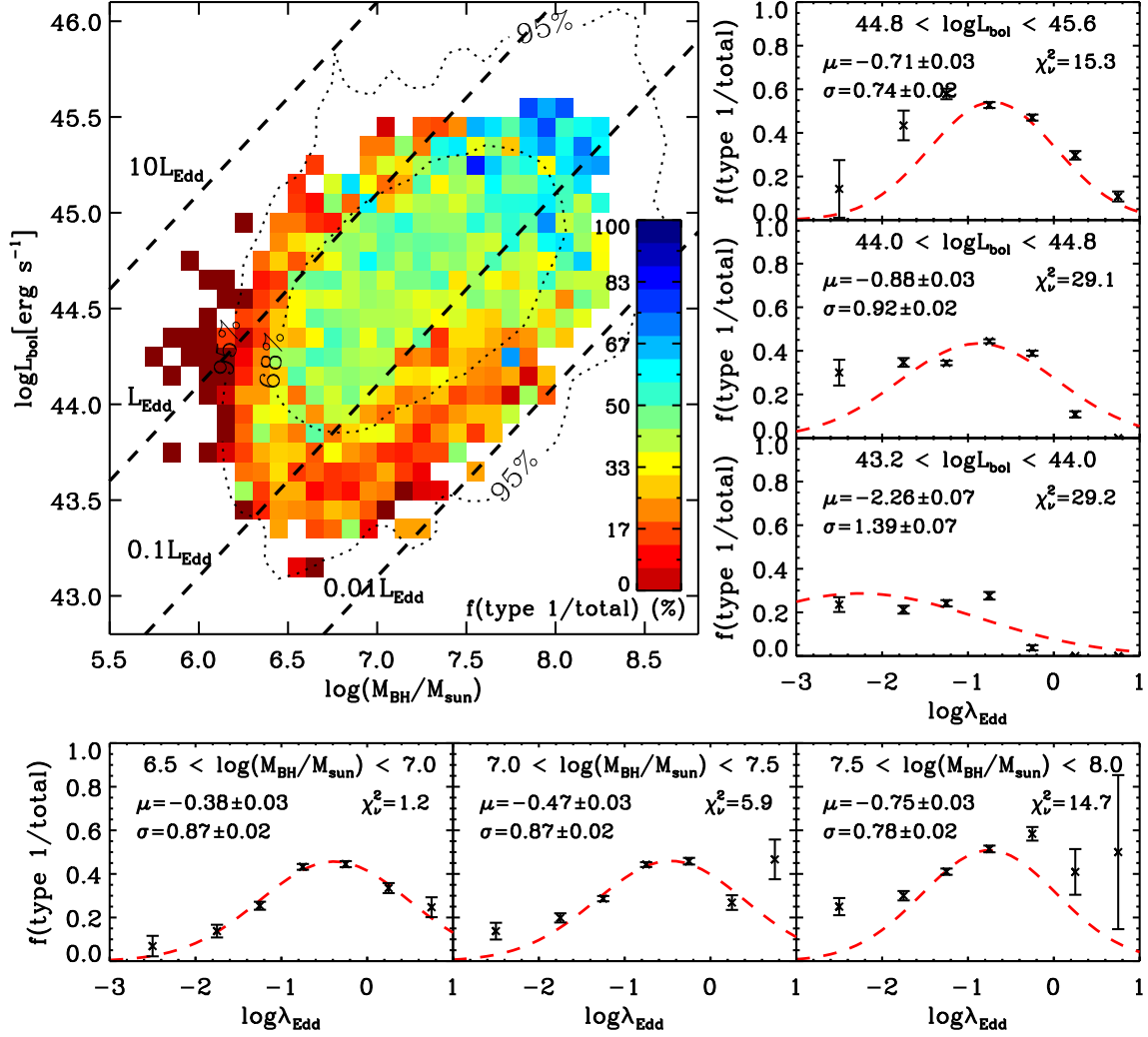


FIG. 16.— Type 1 AGN fraction with respect to bolometric luminosity and black hole mass. The main panel presents the type 1 AGN fraction by using a color grid to denote bins including more than 10 objects in total (type 1 AGNs + type 2 AGNs). Black dotted curves indicate contours (68% and 95%) for type 1 AGNs, and dashed lines indicate various Eddington ratios. (Right panels) Type 1 AGN fraction versus Eddington ratio for various ranges of bolometric luminosity. (Bottom panels) Type 1 AGN fraction versus Eddington ratio for various ranges of black hole mass. Red curves in the right and bottom panels show log-normal fits. Tab. 8 lists all data points, errors, and fitting results.

TABLE 8  
TYPE 1 AGN FRACTION AS A FUNCTION OF EDDINGTON RATIO

$\log \lambda_{\text{Edd}}$	$\log(M_{\text{BH}}/M_{\odot})$			$\log L_{\text{bol}}^{\text{a}}$		
	6.5 – 7.0	7.0 – 7.5	7.5 – 8.0	43.2 – 44.0	44.0 – 44.8	44.8 – 45.6
–3.0 – –2.0	$0.07 \pm 0.05$	$0.14 \pm 0.04$	$0.25 \pm 0.04$	$0.24 \pm 0.03$	$0.30 \pm 0.06$	$0.14 \pm 0.13$
–2.0 – –1.5	$0.14 \pm 0.03$	$0.20 \pm 0.02$	$0.30 \pm 0.02$	$0.21 \pm 0.02$	$0.35 \pm 0.02$	$0.43 \pm 0.07$
–1.5 – –1.0	$0.25 \pm 0.02$	$0.29 \pm 0.01$	$0.41 \pm 0.01$	$0.24 \pm 0.01$	$0.34 \pm 0.01$	$0.58 \pm 0.02$
–1.0 – –0.5	$0.43 \pm 0.01$	$0.44 \pm 0.01$	$0.51 \pm 0.02$	$0.28 \pm 0.02$	$0.44 \pm 0.01$	$0.53 \pm 0.01$
–0.5 – 0.0	$0.44 \pm 0.01$	$0.46 \pm 0.02$	$0.58 \pm 0.03$	$0.04 \pm 0.01$	$0.39 \pm 0.01$	$0.47 \pm 0.01$
0.0 – 0.5	$0.33 \pm 0.02$	$0.27 \pm 0.03$	$0.41 \pm 0.10$	$0.00 \pm 0.00$	$0.11 \pm 0.01$	$0.30 \pm 0.02$
0.5 – 1.0	$0.25 \pm 0.05$	$0.47 \pm 0.09$	$0.50 \pm 0.35$	$0.00 \pm 0.00$	$0.00 \pm 0.00$	$0.11 \pm 0.02$
fit with log-normal distribution						
$\mu$	$-0.38 \pm 0.03$	$-0.47 \pm 0.03$	$-0.75 \pm 0.03$	$-2.26 \pm 0.07$	$-0.88 \pm 0.03$	$-0.71 \pm 0.03$
$\sigma$	$0.87 \pm 0.02$	$0.87 \pm 0.02$	$0.78 \pm 0.02$	$1.39 \pm 0.07$	$0.92 \pm 0.02$	$0.74 \pm 0.02$
$\chi^2_{\nu}$	1.2	5.9	14.7	29.2	29.1	15.3

<sup>a</sup> in unit of  $[\text{erg s}^{-1}]$

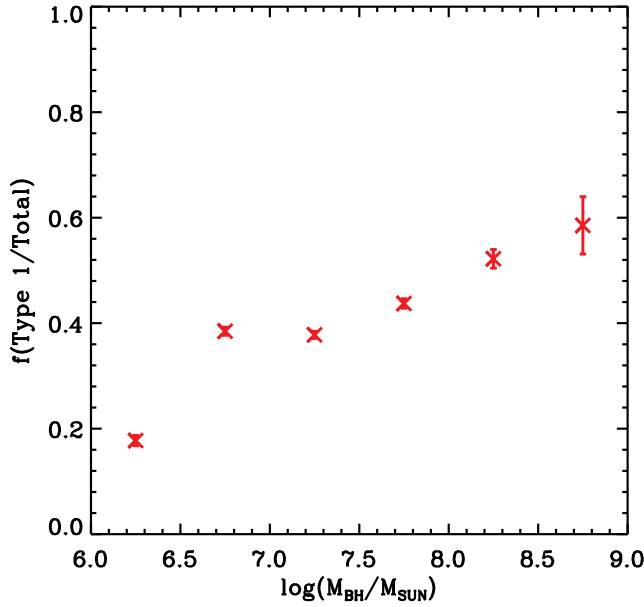


FIG. 17.— Type 1 AGN fraction with respect to black hole mass.

accretion disk. Because outflow velocities are expected to decrease with  $R$ , the torus outflow scenario also predicts luminosity dependency of the type 1 AGN fraction. Again, however, our data shown in Fig. 16 do not seem to show a monotonic luminosity dependence. Rather, the covering factor of the torus seems to vary with the Eddington ratio, showing a favored spot in the black hole mass and bolometric luminosity diagram.

From the perspective of black hole mass, i.e., horizontal axis of Fig. 16, smaller black hole mass regime shows lower type 1 AGN fraction. It can clearly be seen in Fig. 17 that type 1 AGN fraction gradually increases with black hole mass. As lower mass black hole allows more orbits close to the centre, the accretion disk is more

heated up overionising broad  $H\alpha$  emitting BLRs. However, it should be noted that further detailed analysis is required to confirm this interpretation.

It is also noteworthy that there are discussions regarding the effect of the Eddington ratio and luminosity of AGN on the presence of BLRs. Nicastro et al. (2003) presented evidence that hidden BLRs are found above the certain threshold value of Eddington ratio based on their earlier model (Nicastro 2000) and a handful of spectropolarimetric samples with X-ray data. They claimed that BLRs are formed by accretion disk instabilities, and the critical radius becomes smaller than the innermost stable orbit under the low enough accretion rates. Similarly, Elitzur & Ho (2009) showed that type 1 AGNs are not found below a certain luminosity based on disk-wind scenario. Their data supports the disappearance of BLRs at radiatively inefficient accretion regime which is far more lower than our data presented.

In conclusion, our new type 1 AGN catalogue seems to pose a serious challenge to the current favorite versions of the AGN unified model, but the improved database will hopefully help us better understand the AGN classifications in the end.

#### ACKNOWLEDGMENTS

We are grateful to the anonymous referee for a number of clarifications that improved the quality of the manuscript. We also thank Kurt Soto for his valuable comments. KO acknowledges support from Swiss Government Excellence Scholarship for the academic year 2013 – 2014 (ESKAS No.2013.0308). SKY acknowledges support from the National Research Foundation of Korea (Doyak 2014003730). SKY acted as the corresponding author. KS and MK gratefully acknowledge support from Swiss National Science Foundation Professorship grant PP00P2 138979/1 and MK support from SNSF Ambizione grant PZ00P2 154799/1. This study was performed under the DRC collaboration between Yonsei University and the Korea Astronomy and Space Science Institute.

#### REFERENCES

- Abazajian, K., Adelman-McCarthy, J. K., Agüeros, M. A., et al. 2004, *AJ*, **128**, 502
- Alonso-Herrero, A., Pérez-González, P. G., Rieke, G. H., et al. 2008, *ApJ*, **677**, 127
- Antonucci, R. 1993, *ARA&A*, **31**, 473
- Arnaud, K. A. 1996, in *Astronomical Society of the Pacific Conference Series*, Vol. 101, *Astronomical Data Analysis Software and Systems V*, ed. G. H. Jacoby & J. Barnes, 17
- Assef, R. J., Stern, D., Kochanek, C. S., et al. 2013, *ApJ*, **772**, 26
- Baldwin, J. A., Phillips, M. M., & Terlevich, R. 1981, *PASP*, **93**, 5
- Ballo, L., Cristiani, S., Fasano, G., et al. 2007, *ApJ*, **667**, 97
- Bruzual, G., & Charlot, S. 2003, *MNRAS*, **344**, 1000
- Burlon, D., Ajello, M., Greiner, J., et al. 2011, *ApJ*, **728**, 58
- Cappellari, M., & Emsellem, E. 2004, *PASP*, **116**, 138
- Cash, W. 1979, *ApJ*, **228**, 939
- Elitzur, M., & Ho, L. C. 2009, *ApJ*, **701**, L91
- Elitzur, M., & Shlosman, I. 2006, *ApJ*, **648**, L101
- Evans, I. N., Primini, F. A., Glotfelty, K. J., et al. 2010, *ApJS*, **189**, 37
- Fabian, A. C., Rees, M. J., Stella, L., & White, N. E. 1989, *MNRAS*, **238**, 729
- Falcke, H., Gopal-Krishna, & Biermann, P. L. 1995, *A&A*, **298**, 395
- Filippenko, A. V. 1997, *ARA&A*, **35**, 309
- Greene, J. E., & Ho, L. C. 2005, *ApJ*, **630**, 122
- . 2007, *ApJ*, **667**, 131
- Grimes, J. A., Rawlings, S., & Willott, C. J. 2004, *MNRAS*, **349**, 503
- Grimes, J. P., Heckman, T., Strickland, D., & Ptak, A. 2005, *ApJ*, **628**, 187
- Grimm, H.-J., Gilfanov, M., & Sunyaev, R. 2003, *MNRAS*, **339**, 793
- Haardt, F., & Maraschi, L. 1991, *ApJ*, **380**, L51
- Haggard, D., Green, P. J., Anderson, S. F., et al. 2010, *ApJ*, **723**, 1447
- Hao, L., Strauss, M. A., Fan, X., et al. 2005, *AJ*, **129**, 1795
- Hasinger, G. 2004, *Nuclear Physics B Proceedings Supplements*, **132**, 86
- . 2008, *A&A*, **490**, 905
- Heckman, T. M., Kauffmann, G., Brinchmann, J., et al. 2004, *ApJ*, **613**, 109
- Hill, G. J., Goodrich, R. W., & Depoy, D. L. 1996, *ApJ*, **462**, 163
- Hönig, S. F., & Beckert, T. 2007, *MNRAS*, **380**, 1172
- Huchra, J., & Burg, R. 1992, *ApJ*, **393**, 90
- Kauffmann, G., Heckman, T. M., Tremonti, C., et al. 2003, *MNRAS*, **346**, 1055
- Kawaguchi, T., Shimura, T., & Mineshige, S. 2001, *ApJ*, **546**, 966
- Kennicutt, Jr., R. C. 1998, *ApJ*, **498**, 541
- Kewley, L. J., Dopita, M. A., Sutherland, R. S., Heisler, C. A., & Trevena, J. 2001, *ApJ*, **556**, 121
- Krolik, J. H., & Begelman, M. C. 1988, *ApJ*, **329**, 702
- Lawrence, A. 1991, *MNRAS*, **252**, 586
- Lehmer, B. D., Alexander, D. M., Bauer, F. E., et al. 2010, *ApJ*, **724**, 559
- Maia, M. A. G., Machado, R. S., & Willmer, C. N. A. 2003, *AJ*, **126**, 1750
- Maiolino, R., & Rieke, G. H. 1995, *ApJ*, **454**, 95



TABLE 9  
PROPERTIES OF TYPE 1 AGNs<sup>a</sup>

SDSS ObjID	RA <sup>b</sup> [deg]	DEC <sup>b</sup> [deg]	z	logL[O III] [erg s <sup>-1</sup> ]	FWHM <sub>bH<math>\alpha</math></sub> [km s <sup>-1</sup> ]	EW <sub>bH<math>\alpha</math></sub> [Å]	logL <sub>bH<math>\alpha</math></sub> [erg s <sup>-1</sup> ]	log( $M_{\text{BH}}/M_{\odot}$ ) <sup>c</sup>	logL <sub>bol</sub> <sup>d</sup> [erg s <sup>-1</sup> ]	notes <sup>e</sup>
587728309632106568	152.29813	2.74760	0.06319	39.98	6671.6 ± 89.2	56.9	41.18	7.55 <sup>+0.05</sup> <sub>-0.05</sub>	43.52	N
587729229300760654	244.32758	52.48399	0.06319	39.85	3857.1 ± 159.1	18.6	40.77	6.83 <sup>+0.06</sup> <sub>-0.05</sub>	43.40	N
587724648182448168	178.02454	-3.50442	0.06320	41.46	4120.9 ± 23.7	170.6	42.10	7.62 <sup>+0.05</sup> <sub>-0.05</sub>	45.00	3
587732482763915309	229.21861	39.90373	0.06323	40.74	4606.0 ± 167.3	18.9	41.06	7.15 <sup>+0.06</sup> <sub>-0.06</sub>	44.29	N
587742061069074446	175.31734	21.93939	0.06323	42.26	3850.9 ± 17.2	204.5	42.38	7.72 <sup>+0.06</sup> <sub>-0.05</sub>	45.81	3
588295842853159031	156.47778	47.24015	0.06325	40.50	3736.5 ± 95.4	26.7	40.96	6.91 <sup>+0.05</sup> <sub>-0.04</sub>	44.05	XN
587742188840943769	208.48332	21.99855	0.06333	40.39	6186.1 ± 101.2	41.7	41.34	7.57 <sup>+0.06</sup> <sub>-0.06</sub>	43.93	N
587731522275311780	123.89431	35.28634	0.06337	40.75	3480.6 ± 39.1	64.5	41.36	7.07 <sup>+0.06</sup> <sub>-0.04</sub>	44.29	3
587739097528270899	190.37257	37.36722	0.06340	41.26	4243.6 ± 32.3	118.5	41.84	7.50 <sup>+0.05</sup> <sub>-0.05</sub>	44.81	3
588023668097875996	183.71655	19.01166	0.06340	40.37	4734.6 ± 75.3	57.1	41.19	7.25 <sup>+0.05</sup> <sub>-0.04</sub>	43.91	3
587739827668844624	224.67809	21.60277	0.06343	38.78	2343.2 ± 9.9	243.0	42.18	7.16 <sup>+0.04</sup> <sub>-0.04</sub>	42.32	3
587730847966888351	326.21548	0.47415	0.06344	40.62	3770.0 ± 83.4	34.5	40.89	6.88 <sup>+0.04</sup> <sub>-0.04</sub>	44.17	G
587742060531089515	172.65646	21.23771	0.06347	40.14	4781.4 ± 149.6	24.6	40.99	7.14 <sup>+0.06</sup> <sub>-0.05</sub>	43.68	G
587741720679612540	203.14915	25.29781	0.06353	39.93	4987.1 ± 99.6	52.1	41.23	7.32 <sup>+0.05</sup> <sub>-0.05</sub>	43.48	N
587734690883698778	127.78179	5.35165	0.06356	40.93	4447.3 ± 65.3	55.8	41.34	7.27 <sup>+0.05</sup> <sub>-0.05</sub>	44.48	3G
587739721376792749	225.65990	24.22051	0.06361	41.11	4392.0 ± 43.7	92.6	41.42	7.31 <sup>+0.04</sup> <sub>-0.04</sub>	44.66	3G
587741600421314588	184.55672	26.40504	0.06361	40.21	2131.0 ± 20.9	70.4	41.17	6.52 <sup>+0.02</sup> <sub>-0.02</sub>	43.75	3G
588017977836765242	232.13504	28.96422	0.06362	40.72	4744.4 ± 64.9	61.1	41.42	7.38 <sup>+0.05</sup> <sub>-0.05</sub>	44.27	G
587739810484650051	208.23201	25.48323	0.06366	41.30	2505.1 ± 14.0	175.1	41.84	7.03 <sup>+0.03</sup> <sub>-0.03</sub>	44.84	3HG

NOTE. — (This table is available in its entirety in a machine-readable form in the online journal. A portion is shown here for guidance regarding its form and content.)

<sup>a</sup> This table is also available in the OSSY web site (<http://www.gem.yonsei.ac.kr/ossy/>)

<sup>b</sup> J2000 coordinates

<sup>c</sup> Black hole mass derived from the method developed by [Greene & Ho \(2005\)](#)

<sup>d</sup> Bolometric luminosity derived from the method developed by [Heckman et al. \(2004\)](#)

<sup>e</sup> Flag presenting the newly identified type 1 AGNs ('N') with the *Chandra* X-ray data ('X'). '3' indicates the SDSS *SpecClass* 3 objects. 'H' and 'G' indicate the object also identified by [Hao et al. \(2005\)](#) and [Greene & Ho \(2007\)](#), respectively.

Markwardt, C. B. 2009, in *Astronomical Society of the Pacific Conference Series*, Vol. 411, *Astronomical Data Analysis Software and Systems XVIII*, ed. D. A. Bohlender, D. Durand, & P. Dowler, 251

Mineo, S., Gilfanov, M., & Sunyaev, R. 2011, *Astronomische Nachrichten*, **332**, 349

Nicastro, F. 2000, *ApJ*, **530**, L65

Nicastro, F., Martocchia, A., & Matt, G. 2003, *ApJ*, **589**, L13

Oh, K., Sarzi, M., Schawinski, K., & Yi, S. K. 2011, *ApJS*, **195**, 13

Osterbrock, D. E., & Shaw, R. A. 1988, *ApJ*, **327**, 89

Peng, C. Y., Ho, L. C., Impey, C. D., & Rix, H.-W. 2002, *AJ*, **124**, 266

Ranalli, P., Comastri, A., & Setti, G. 2003, *A&A*, **399**, 39

Salim, S., Rich, R. M., Charlot, S., et al. 2007, *ApJS*, **173**, 267

Sánchez, S. F., Jahnke, K., Wisotzki, L., et al. 2004, *ApJ*, **614**, 586

Sánchez-Blázquez, P., Peletier, R. F., Jiménez-Vicente, J., et al. 2006, *MNRAS*, **371**, 703

Sarzi, M., Falcón-Barroso, J., Davies, R. L., et al. 2006, *MNRAS*, **366**, 1151

Schawinski, K., Thomas, D., Sarzi, M., et al. 2007, *MNRAS*, **382**, 1415

Schawinski, K., Treister, E., Urry, C. M., et al. 2011, *ApJ*, **727**, L31

Schawinski, K., Virani, S., Simmons, B., et al. 2009, *ApJ*, **692**, L19

Schawinski, K., Urry, C. M., Virani, S., et al. 2010, *ApJ*, **711**, 284

Schneider, D. P., Richards, G. T., Hall, P. B., et al. 2010, *AJ*, **139**, 2360

Simmons, B. D., & Urry, C. M. 2008, *ApJ*, **683**, 644

Simmons, B. D., Lintott, C., Schawinski, K., et al. 2013, *MNRAS*, **429**, 2199

Simpson, C. 2005, *MNRAS*, **360**, 565

Steffen, A. T., Barger, A. J., Cowie, L. L., Mushotzky, R. F., & Yang, Y. 2003, *ApJ*, **596**, L23

Stern, J., & Laor, A. 2012, *MNRAS*, **423**, 600

Ueda, Y., Akiyama, M., Ohta, K., & Miyaji, T. 2003, *ApJ*, **598**, 886

Urry, C. M., & Padovani, P. 1995, *PASP*, **107**, 803

Vanden Berk, D. E., Shen, J., Yip, C.-W., et al. 2006, *AJ*, **131**, 84

Wachter, K., Leach, R., & Kellogg, E. 1979, *ApJ*, **230**, 274

Westoby, P. B., Mundell, C. G., & Baldry, I. K. 2007, *MNRAS*, **382**, 1541

Willott, C. J., Rawlings, S., Blundell, K. M., & Lacy, M. 2000, *MNRAS*, **316**, 449

York, D. G., Adelman, J., Anderson, Jr., J. E., et al. 2000, *AJ*, **120**, 1579

Zdziarski, A. A., Poutanen, J., & Johnson, W. N. 2000, *ApJ*, **542**, 703